Social Interface and Interaction Design for Group Recommender Systems

THÈSE N^{O} 6353 (2015)

PRÉSENTÉE LE 30 JANVIER 2015

À LA FACULTÉ INFORMATIQUE ET COMMUNICATIONS

GROUPE SCI IC PFP

PROGRAMME DOCTORAL EN INFORMATIQUE, COMMUNICATIONS ET INFORMATION

ÉCOLE POLYTECHNIQUE FÉDÉRALE DE LAUSANNE

POUR L'OBTENTION DU GRADE DE DOCTEUR ÈS SCIENCES

PAR

Yu CHEN

acceptée sur proposition du jury:

Prof. E. Telatar, président du jury
Dr P. Pu Faltings, directrice de thèse
Prof. A. Geissbuhler, rapporteur
Dr D. Gillet, rapporteur
Dr D. Lalanne, rapporteur

ÉCOLE POLYTECHNIQUE
FÉDÉRALE DE LAUSANNE

Suisse
2015

While drinking water,
do not forget those who dug the well.
— Chinese proverb

To my parents and all my teachers

Acknowledgements

Towards finishing this thesis, I would like to express my deepest gratitude to my advisor Dr. Pearl Pu. Without her effort, I could never have encountered such an interesting and challenging research project. When I look back at the four years in HCI Group, I am amazed how she has helped me identify my strength and uncover my potentiality. She skillfully guided me to master all essential skills in multi-disciplines as an HCI researcher. My sincere appreciation for her encouragement when I was down and proposing more challenging goals when I conquered the previous one. I feel so fortunate to have her as my advisor! My sincere gratitude also goes to Swiss National Science Foundation for generously sponsoring this exciting research project for the four years.

I would like to thank my jury members – Prof. Emre Telatar, Dr. Denis Gillet, Prof. Antoine Geissbuhler, and Dr. Denis Lalanne – for their time and support in the thesis. Thanks a lot for accommodating on scheduling the defense!

I feel so lucky to have a group of wonderful colleagues and excellent researchers. First of all, I want to thank Rong for providing feedback and accompanying me for almost my entire Ph.D. journey. She is a great role model with her persistent research attitudes. I appreciate Jiyong for sharing experience and offering feedback for my research. I enjoy the collaboration and in-depth yet joyful conversations with Onur, who is always filled with positive energy. It is a great experience working with Valentina and Lionel on teaching activities and sharing research ideas. Many thanks go to our secretaries Ms. Sylvie Thomet and Ms. Madeleine Robert for their kind support. All of them help me build up a beautiful memory in HCI Group.

I am thankful for my Master thesis advisor Dr. Zheng Yan, who had been constantly encouraging and supporting me to pursue an academic career since I first met her five years ago.

Thank Lan Huang and Yu Su for jogging on this Ph.D. marathon with me. Interacting with them over the continents across different time zones greatly inspired my research.

I planned to name all my teachers and friends both in Switzerland and across the oceans, who have been kindly cheering me up during the four years. But the list starts to become infinite, and I fall short of words. I also owe a lot to my relatives in China who have been providing so much care for my family when I was absent.

Last but not least, thank my papa and mama for their unconditional love and providing me all I need to pursue my dreams. I love you!

Lausanne, 8 Aug 2014

Abstract

Group recommender systems suggest items of interest to a group of people. Traditionally, group recommenders provide recommendations by aggregation the group members' preferences. Nowadays, there is a trend of decentralized group recommendation process that leverages the group dynamics and reaches the recommendation goal by allowing group members to influence and persuade each other. So far, the research on group recommender systems mainly focuses on the how to optimize the preference aggregation and enhance the accuracy of recommendations. There is a lack of emphasis on the users' social experience, such as interpersonal relationship, emotion exchange, group dynamics, etc. We define the space where user-user interaction occurs in social software as ***social interfaces***. In this thesis, we aim to design and evaluate social interfaces and interactions for group recommender systems.

We start with surveying the state-of-the-art of user issues in group recommender systems and interface and interaction design in the broad sense of social applications. We present ten applications and their evaluation via user studies, which lead to a preliminary set of social interface and interaction design guidelines. Based on these guidelines, we develop group recommender systems to investigate the design issues.

We then study *social interfaces* for group recommender systems. We present the design and development of an experimental platform called ***GroupFun*** that recommends music to a group of users. We then study the impact of emotion awareness in group recommender systems. More concretely, we design and implement two different methods for emotion awareness: ***CoFeel*** and ***ACTI*** that visualize emotions using color wheels, and ***empatheticons*** that present emotions using dynamic animations of users' profile pictures. Our user studies show that emotion awareness tools can help users familiarize with other members' preferences, enhance their interpersonal relationships, increase the sense of connectedness in distributed social interactions, and result in higher consensus and satisfaction in group recommendations.

We also examine *social interactions* for persuasive technologies. We design and develop a mobile social game called ***HealthyTogether*** that enables dyads to exercise together. With this platform, we study how different social interaction mechanisms, such as social accountability, competition, cooperation, and team spirits, can help users motivate and influence each other in physical exercises. We conducted three user studies lasting for up to ten weeks with a total of 80 users. This is so far the first series of studies that extensively investigate the social incentives for behavior change. Results show that HealthyTogether can effectively stimulate dyads to walk compared with when they exercise alone. Additionally, being accountable for each other's performance enhances interpersonal relationships. Supporting users to cooperate

on health goals significantly improve their number of steps. When designing competition in the applications, it is crucial to help users to choose comparable buddies. Finally, teamwork in exercises not only helps users to increase their steps, but also help them sustain in exercise.

Furthermore, we present an evaluation framework for social persuasive applications. The framework aims at modeling how social strategies and social influence affect users' attitudes and behavioral intentions towards the system. We also report the results of a pilot study to validate the framework using data collected from a survey with 46 entries.

Finally, we derive a set of guidelines for social interface and interaction design for group recommender systems. The guidelines cover six crucial requirements that influence users' social experiences: mutual awareness, intimacy, immediacy, group dynamics, motivation, and playfulness. The guidelines can help researchers and practitioners effectively design social experiences for not only group recommenders but also other social software and identify important areas in which to invest development resources.

Key words: social interface, social interaction, design guidelines, group recommender systems, emotion awareness, healthy behavior change, empirical studies, evaluation framework

Résumé

Un système de recommandation de groupe suggère des objets qui pourraient intéresser un certain groupe de personne. Traditionnellement, ces systèmes fournissent des recommandations en tenant compte des préférences de ses membres. Actuellement, la tendance est de décentraliser le processus de recommandation de groupe afin de maximiser la dynamique de groupe ainsi que d'atteindre l'objectif du système, ceci, en permettant aux membres du groupe de s'influencer et de se persuader mutuellement. Jusqu'à maintenant, la recherche sur les systèmes de recommandation de groupe se focalisait principalement sur le moyen d'optimiser l'agrégation de préférences et d'améliorer la précision des recommandations. Très peu de recherches mettent l'accent sur l'expérience sociale de l'utilisateur du système de recommandation de groupe tels que les relations interpersonnelles, les échanges émotionnels, la dynamique du groupe, etc. Nous définissons un espace où l'interaction entre les utilisateurs se manifeste à travers des logiciels sociaux tels que les ***interfaces sociales***. Dans cette thèse, nous avons pour objectif de designer et d'évaluer les interfaces sociales ainsi que les interactions au sein d'un système de recommandation de groupe.

En premier lieu, nous commencerons par analyser l'état de l'art relié aux difficultés rencontrées par les utilisateurs au sein des systèmes de recommandation de groupe et des applications sociales, aussi bien par rapport au design de l'interface que de l'interaction en générale. A travers une étude de recherche sur les utilisateurs et un processus d'évaluation approfondi, nous présenterons dix applications. Ceci nous permettra de déduire un ensemble d'instructions préliminaires pour la conception de l'interface et de l'interaction d'un tel système. Nous développerons ensuite un système de recommandation de groupe qui se basera sur ces directives afin d'examiner plus en détail les problématiques reliées au design de tels outils.

En second lieu, nous étudierons les interfaces sociales des systèmes de recommandation de groupe. Nous commencerons par présenter une plateforme expérimentale appelée ***GroupFun*** qui recommande de la musique à un groupe d'utilisateurs. Ensuite, nous analyserons l'impact de la sensibilité émotionnelle dans les systèmes de recommandation de groupe. Plus concrètement, nous concevrons et implémenterons deux méthodes différentes pour la sensibilité émotionnelle : ***ACTI*** et ***CoFeel*** qui visualisent les émotions en utilisant des roues colorées et ***empatheticons*** qui présente les émotions en animant dynamiquement les photos de profil des utilisateurs selon l'émotion. Nos études d'utilisateurs montrent que les outils de sensibilisation émotionnelle aident les utilisateurs à se familiariser avec les préférences des autres membres du groupe, et améliorent ainsi les relations interpersonnelles dans le groupe,

le sentiment d'appartenance dans les interactions sociales distribuées et le consensus obtenu. Au même titre, nous avons remarqué une plus grande satisfaction dans les recommandations de groupe.

Ensuite, nous examinerons l'impact des interactions sociales dans le cadre d'applications de remise en forme. Nous présenterons ***HealthyTogether***, une application mobile de remise en forme conçue pour permettre à des binômes de s'entrainer ensemble. Avec cette plateforme, nous étudierons comment différents mécanismes d'interactions sociales, comme la compatibilité sociale, la compétition, la coopération et l'esprit d'équipe, motivent à faire de l'exercice physique. Nous avons conduit trois études sur les utilisateurs qui ont duré jusqu'à 10 semaines avec un total de 80 utilisateurs. C'est, à ce jour, la seule série d'étude ayant étudié extensivement les récompenses sociales incitant au changement de comportement. Les résultats que nous avons obtenus montrent que HealthyTogether peut efficacement accroître la marche chez les binômes en comparaison avec des stratégies d'exercice individuel. En outre, avoir un impact sur la performance de l'autre améliore également les relations interpersonnelles. Encourager l'autre utilisateur à coopérer pour atteindre des objectifs communs en rapport avec la santé améliore significativement le nombre de pas mesurés. Pour toute compétition au sein des applications, il est crucial d'aider les utilisateurs à choisir un binôme similaire. Finalement, le travail d'équipe dans l'approche de l'exercice physique aide non seulement les utilisateurs à améliorer leur nombre de pas, mais également à maintenir la pratique de l'exercice.

Pour poursuivre, nous présenterons de manière détaillée un système d'évaluation des applications sociales persuasives dans le contexte de la remise en forme généralisée. Ce système vise à modéliser l'impact des différentes stratégies et influences sociales sur l'attitude et les intentions de l'utilisateur vis à vis de l'application. Nous rapporterons également les résultats d'une étude pilote pour valider cet ensemble d'outils en utilisant les données collectées à partir d'un sondage avec 46 entrées.

Finalement, nous dériverons un ensemble de règles de conception pour le design de l'interface et de l'interaction d'un système de recommandation de groupe. Ces directives couvrent six critères cruciaux influençant l'expérience sociale des utilisateurs : la conscience réciproque, l'intimité, l'instantanéité, les dynamiques de groupe, la motivation et l'aspect ludique. Ces règles peuvent aider les chercheurs et les développeurs à concevoir de manière plus efficace des interfaces non seulement de systèmes de recommandation de groupe mais également d'autres applications sociales. Ils seront à même d'identifier les domaines pertinents dans lesquels investir les ressources pour leur développement.

Mots clefs : interfaces sociales, interactions sociales, règles de conception, études d'utilisation, directives, systèmes de recommandation de groupe, sensibilisation émotionnelle, changement de comportement sains, système d'évaluation

Contents

List of Figures

List of Tables

1 Introduction

The past decades have witnessed the reshaping of online services from content-based to user-based. Search engines no longer provide universal results to all users with the same keywords, but adapt searching results to users' profiles, such as location, browsing history and social networks. The Google Search results from users in Switzerland are different from users in the US. Results that are previously shared by users' social networks are also shown on top of the results list. E-commerce websites do not merely display the products that users try to buy, but actively recommend items that users may like based on their preferences or purchase history. Amazon is a typical example of e-commerce website that intelligently suggests to users "what you may like." Online music services also evolve from allowing users to actively search for music they want to listen to suggesting personalized music lists. They aim at continually providing music that meet users' tastes based on their listening or rating behavior. Designing around users instead of content has become a trend in the community of human-computer interaction.

Building around users is more than providing services that personalize their individual needs. Human beings are social creatures, and this calls for the attention of building interfaces and services around people's social needs. Humans have the internal needs to connect with other people either online or offline [18]. The booming of social media demonstrates how building services and systems around users' social needs captivates users. Social network services such as Facebook and Twitter enable users to engage others actively online by tagging, following, re-tweeting, etc. Recent studies show that the more users use social media, the happier they are [89]. The gaming industry also presents the success by integrating users' social needs. The social games from Zynga [16], such as Farmville [4], Cityville [1], DrawSomething [3], may not excel other games in terms of game content or graphical interfaces, but they are designed around users' social ties.

Small groups of friends are identified as the key to influence on the social web [18]. Such groups can be constituted by families selecting a recipe together, colleagues working on the same project, and social club members planning a culture event. In group environments, decision-making therefore becomes an issue due to information overload. ***Group recom-***

mender systems (GRSs) aim to alleviate information overload by suggesting items to a group of people based on group members' preferences [35]. In the recent years, group recommendation process has been extended to a "distributed" way by leveraging the group dynamics and reaching the recommendation goal by allowing group members to influence and persuade each other. Typical examples include persuasive technologies that involve a group of users to accomplish health goals through social persuasion.

Group recommendation problem is not only "the sum of members [74]." As the audiences move from individuals to groups of people, challenges arise such as aggregating preferences and arriving at equilibrium point of expectations. More importantly, users' attitudes of group recommender systems are also influenced by their interactions with other group members. We define the elements of a system that mediate the interaction between multiple members online as ***social interfaces***. Such elements include graphical interfaces that visualize all group members' information, embedded game rules that enforce users to interact online, communication components, etc. To the best of our knowledge, research that comprehensively explores interface and interaction design in group recommender systems is lacking. This thesis aims at investigating these issues and providing design guidelines.

This thesis offers four main contributions in the human-computer interaction (HCI) community. First, it theoretically formalizes and articulates the importance of building systems and applications around users' social needs and interactions. Second, it presents the design, evaluation of emotion awareness interfaces for group recommenders to enhance users' intimacy and immediacy. Third, it investigates how to effectively design social incentives to motivate and engage users in such systems. Last but not least, it presents design guidelines of social interface and interaction for future researchers and practitioners to develop systems that build around users' social needs. Part of this chapter is published in [35].

1.1 Research Motivation

The concept of "social interface" is not new. Current practices mainly focus on designing interfaces when users have to collaboratively work in a social environment, also known as groupware. Typical examples include online group working tools such as Google Drive and Dropbox, online communication tools such as Skype. These systems provide effective solutions that allow users to complete traditionally-collocated tasks in a distributed way. In the domain of group recommender systems, social issues have been translated into interfaces that aim at enhancing the mutual awareness among group members, such membership awareness, preference awareness, decision awareness, etc. In this thesis, we are interested in designing novel interfaces and interaction mechanisms for group recommenders at another level – engaging users and improving the relationship among group members to reach their goals.

Many group recommender systems are also designed to leverage the social influence and persuasion among group members to fulfill their goals. Typical examples include persuasive technologies for behavior change, such as doing more physical activities, taking more balanced diet, improving sleep quality, maintaining a happier mood, etc. Ideally, users could accomplish the tasks independently. However, research shows that users tend to fail their goals for healthier behavior change [80] even though they are fully aware of the benefits and necessity. This is mainly due to lack of time and lack of urgency of such tasks. Recent studies have shown the effectiveness of applying social persuasion and recommendation in motivating healthier lifestyle. Thus, this thesis examines how to design social interaction schemes to engage users and help them accomplish their shared health goals.

1.2 Research Questions

This thesis presents research challenges in designing social interfaces and interaction schemes for **small groups** in two contexts. The first context covers group recommenders that proactively make suggestions to the groups. We study the interface design using a music group recommender system. The second context covers persuasive technologies that allow users to voluntarily persuade and influence each other. We investigate how to design social interactions in persuasive technologies for physical activities. This thesis thus answers the following research questions.

How to Design Social Interfaces for Group Recommender Systems?

Conventional group recommenders enhance user experience mainly by designing interfaces that increase system performance and recommendation accuracy. However, numerous researchers have shown that performance is not the only criteria that determines users' satisfaction in collaboration [89]. Some research shows that the closer relationship the group members have, the more fluent their collaboration is likely to be [90]. Other research also shows that users tend to have higher satisfaction in collaborative work when the group members are more alike each other. So, how to design social interfaces to enhance users' satisfaction in accomplishing collective goals in group recommenders? More concretely, we answer the following questions in the thesis.

• What are the current practices in designing social interfaces in group recommender systems? This thesis presents the state-of-the-art of user issues in group recommender systems. The main techniques in enhancing user experience in group recommender systems are translated into designing interfaces that enhance mutual awareness among group members. So far, mutual awareness comprises membership awareness, preference awareness and decision awareness. We implemented a group recommender system called *GroupFun* that meets the above criteria.

• What other design dimensions should be considered in designing social interfaces for group recommender systems? Researchers have identified the role of emotions in making group decisions and that users' satisfaction towards recommended items can be modeled using their affective states. Additionally, users' affective states and their decisions can be influenced by each other in a group. Thus, we designed emotion awareness interfaces to stimulate positive group influence.

• How to create emotion awareness tools in group recommender systems? There is a trade-off between designing emotion-aware interfaces that may stimulate positive group influence and ask users to provide emotion information that incurs additional user effort. We created three emotion-aware interfaces that meet the requirements of accuracy and engagement. The first two interfaces visualize emotions using color wheels and the third one using kinetic images.

• How to evaluate the effectiveness and roles of emotion awareness tools in group recommender systems? We conducted a controlled study to investigate how users perceive the intimacy and immediacy of emotion-aware interfaces in group recommender systems. More importantly, we investigated how emotion awareness features enable positive group influence and user satisfaction in group recommender systems.

How to Design Social Interactions for Group Recommender Systems?

Group recommenders are not limited to having the system provide suggestions. Users in a group can also consciously or unconsciously influence and interact with each other. But how to design such social influence and interactions? We choose the domain of persuasive technologies for behavior change to investigate these questions. Behavior change is a complex, long-term process with high relapse rates [80], although users are usually aware of the significance of behavior change for their health benefits. With the advance of pervasive sensing and wearable sensors, there are emerging solutions in the market and the academic research. These sensors help users to track their activities and enhance their awareness of their habits and lifestyles. However, the solutions are prone to the novelty effect, meaning users tend to fall back into old habits after a certain period. Recent findings have provided substantial evidence that social influence can effectively motivate users to be more active and sustain in the process of healthy behavior change. In this thesis, we focus on persuasive technologies for physical activities to study behavior change. More concretely, this thesis aims at answering the following research questions.

• What are the design requirements for persuasive technologies for behavior change? We surveyed the current literature on designing social influence for pervasive fitness applications. Based on the current practices, we have developed a mobile game called *HealthyTogether* that allows users to exercise together using gamification methods. The design of HealthyTogether has evolved iteratively during three user studies.

• Can social interaction schemes motivate users to do more physical activities? We conducted three user studies with 80 users for up to ten weeks to compare users' activities when they are exercising alone and using HealthyTogether. We used qualitative methods such as user interviews and diaries to gain in-depth understanding on users' experience. We also applied quantitative analysis to compare users' physical activities in individual and social settings.

• What social interaction schemes should be deployed to best motivate users? Current literature suggests three major social incentives in persuasive applications: competition, cooperation and community leaderboard. Each mechanism has some merits and suffers from limitations. We used HealthyTogether as an experimental platform and conducted three controlled user studies to compare the various social incentives. This thesis reports the findings on how users behave in each of the social incentives. Are various social interaction schemes useful? If yes, in what condition? If no, how to improve them to maximize the motivation effects?

• How to create social interaction schemes to motivate behavior change? Based on the results of the three longitudinal studies, we propose design guidelines for social interfaces. These guidelines will help future researchers and practitioners to create social incentives to motivate users in their behavior change.

1.3 Main Contributions

This thesis offers both empirical and theoretical contributions in HCI field. Empirically, we have designed three emotion awareness tools for group recommender systems, developed and evaluated them in the case of a music group recommender; we also designed and implemented social interaction schemes that aim to motivate users' behavior change by influencing and persuading each other. We experimentally demonstrated how social interaction schemes can enhance users' group experience, such as their motivation, performance, social relationship, etc. Theoretically, we have proposed an evaluation framework that models social interaction design, users' attitudes and behavioral intentions in persuasive technologies. More concretely, it suggests that designing appropriate social interfaces and interactions augments user-perceived social influence and benefits. These attributes are associated with users' perception of attitudes and their usage intention of the applications. Furthermore, we derive fifteen design guidelines for social interfaces and interactions and highlight their roles in enhancing mutual awareness, intimacy, immediacy, group dynamics, motivation, and playfulness. The evaluation framework and design guidelines can be useful for both researchers and practitioners to identify the previously neglected design elements and better allocate development resources.

1.4 Overview of Dissertation

The rest of the thesis is organized as follows. Chapter 2 surveys the user issues, social interfaces and interaction design in group recommenders. It also presents typical applications that address the social experience and a preliminary set of design guidelines. Chapter 3 presents a music group recommender system called GroupFun as an experimental platform to study social interfaces. This is followed by the design of three emotion awareness tools: ACTI, CoFeel and empatheticons. A user study demonstrates the roles of providing emotion awareness in group recommender systems. Chapter 4 introduces a social game called HealthyTogether that enables dyads to walk together. It is used as an experimental platform for three user studies that extensively investigate various social interactions for group recommenders for behavior change. Chapter 5 proposes an evaluation framework for persuasive technologies for behavior change. It aims to uncover how designing various social strategies associate with users' attitudes and behavioral intentions of the system. Chapter 6 offers a list of guidelines that are derived from the literature and the outcomes of our design and empirical studies. Finally, Chapter 7 concludes the thesis and lays out future directions.

2 State of the Art

In this chapter, we first survey the current literature about interface and interaction design for group recommender systems. We then present ten most representative systems that address interfaces and interaction schemes in social applications. We introduce the design and evaluation of these systems to identify the areas that have received less attention in HCI field and then derive a preliminary set of guidelines for social interface and interaction design. Part of this chapter is published in [31, 30, 35].

2.1 Social Interfaces

In the early times, social interfaces are mainly designed for groupware. Groupware is first defined by Johansen in 1988 as "a computer-based system plus the social group process." [76] Eillis et al. [54] further developed the definition and referred to groupware as "computer-based systems that support groups of people engage in a common task (or goal) and that provide an interface to a shared environment." They stressed the notions of a common task and a shared environment as crucial elements of the definition. They proposed that there should be no rigid classification of groupware and non-groupware. They further classified groupware systems to real-time groupware and non-real-time groupware. They identified five perspectives of successful groupware: distributed systems, communication, human-computer interaction, artificial intelligence and social theory. Grudin also raised a number of challenges for developers for groupware and social dynamics [62]. Early work on designing social interfaces started with investigating computer supported collaborative work for *small groups.*

With the booming of online communities, social networking and social media, social interfaces have expanded to systems for the large social networks. Crumlish and Malone defined social interface as "social user experience interface" in their book *Designing social interfaces* [44]. According to the work, online social experience mainly covers community, social media and social networks. Social interface are created to offer social experiences in the above services. The book summarizes two purposes for social interfaces: social communication and collaboration. It also offers design patterns for social media large online communities. Social patterns, as defined in the book, are the components and pieces of interactivity that are the building blocks of social experiences. These social patterns, derived from major social websites, lead to design guidelines on social interfaces for large online communities.

However, the above research did not address the dynamics and human-human interaction beyond the systems. In other words, the social interfaces deal with how users interact with computers but less address their interaction with people over the computers. As Pentland remarked, "computers are socially ignorant" [102]. He brought forth the notion of socially aware computation and communication that aims to build machines that understand social signaling and context. The purpose of designing socially aware systems is not only to help a group of users complete a task, but enable them to work *together* more smoothly and productively.

2.2 Interface Design for Group Recommender Systems

Group recommenders are typical examples of groupware. Jameson has studied some of the key user issues for group recommender systems [74] and investigated several measures for promoting collaborating and coordination. These measures mainly aim at designing user interfaces to enhance mutual awareness.

Mutual awareness enables members to be aware of the presence of the others so that they share mutual knowledge and can interact with each other. Awareness is a common base for a community to improve contact among group members [110]. In group recommender systems, mutual awareness is translated into a number of interactivity and interface features that have been designed to visualize member presence and highlight the preferences of other members and opinions of the group as a whole [93]. Based on current literature, mutual awareness in group recommender systems includes membership awareness, preference awareness and decision awareness.

Membership awareness allows users to check which users are in the group. Being aware of members in a group facilitates users to decide how to behave and thus enhances trust in a group recommender. Group recommender systems can display active users on system user interface. TV4M [126] enables multiple group members to log into the system in one common user interface and all users were aware who were in the group. In PolyLens [99], all group members can view the membership list and remove themselves from the group. Pocket RestaurantFinder automatically informs users of the joining of a new member (McCarthy 2002). Group membership awareness can also be implemented in the physical world. Collaborative Advisory Travel System (CATS) [93] enables users to cooperate around a common workspace and discuss about a topic in a collocated and synchronized way. The system uses an interactive tabletop which assists up to four users to simultaneously engage in a parallel recommendation sessions.

Preference awareness enables users to be aware of the preferences of other members. A user study on PolyLens [99] reveals that users would like to see each other's preference information, even at the expense of some degree of privacy loss. Thus far, we categorize preference awareness in group recommender systems into three levels: *zero awareness*, *partial awareness*, and *full awareness*.

Zero preference awareness means that users only know their own preferences during preference specification. Typical examples include MusicFX [92], Flytrap [43] and In-Vehicle music recommender [127]. Systems of zero preference awareness are simple to implement; but since users are not able to refer to other members' opinions and thus are less favorable for users to collectively make decisions and trust the systems.

Partial awareness in group recommenders allows users to apply aggregated preference information from other group members. Such systems do not reveal complete preference of individual group members. This helps reducing users' initial effort by allowing members to have the same preference model as certain group members or a whole group [26]. However, users are prone to social loafing, a phenomenon when people contribute less in a social environment than when they work individually.

In full preference awareness, users are fully aware of each other's preferences in a group. One typical technique for full preference awareness is Collaborative Preference Specification (CPS), as presented in CATS [93], Pocket RestaurantFinder [91] and TravelDecisionForum [74]. CPS [74] in group recommender systems has two functions. First, it enables users to persuade other members to specify similar preferences by giving them information that they previously lacked. Second, it supports explaining and justifying a member's preference even if the argument is not generalizable to other members (e.g., "I can't go hiking, because of an injury"). Preference aggregation with full awareness takes into account attitudes and anticipated behaviors of other members.It also encourages assimilation to facilitate conflict minimization.

So far, preference awareness in group recommender systems tends to cause "the rich become richer and the poor become poorer" problems. This is a problem when only a few major users contribute to preference specification while other users are lazy and do not contribute to group activities. The research work on incentives for encouraging the majority of users to contribute to collaborative preference specification is still insufficient. In this thesis, we aim to design novel interfaces and incentives that encourage user contribution.

Decision awareness is important in helping users arrive at a final decision. Decision awareness is a status in which users are aware of the group decision-making process. Existing group recommender systems include the following decision making styles: 1) zero awareness - simply translating the most highly rated solution into action without the consent of any user (e.g. in MusicFX [92], 2) partial awareness - one or a selected set of representatives of the group are responsible for making the final decisions (e.g. in INTRIGUE [23] and PolyLens [99], and 3) full awareness - arriving at final decision through collected discussion, such as face-to-face discussions of all members (e.g., as presented in CATS [93] and/or mediated discussions (e.g., as is implemented in MIAU [82] and Travel Decision Forum [74]).

Name of System	Domain	Membership awareness	Preference awareness	Decision awareness
MusicFX [92]	Music	No	No	No
PolyLens [99]	Movies	Yes	No	No
Flytrap [43]	Music	Yes	No	No
Pocket Restau-rantFinder [91]	Restaurants	No	No	No
MIAU [82]	Purchase	Yes	Yes	Computer-assisted
INTRIGUE [23]	Tourism	No	No	No
Travel Decision Forum [74]	Tourism	Yes	Yes	Agent-assisted
AdaptiveRadio [74]	Music	Yes	No	No
In-Vehicle [127]	Music	No	No	No
TV4M [126]	TV programs	Yes	Yes	No
CATS [93]	Tourism	Yes	Yes	Face-to-face

Table 2.1: Mutual awareness implementation in existing group recommender systems.

From the analysis above, we list implemented features about mutual awareness in existing group recommender systems below. From Table 2.1, it is clear that group recommender systems are paying more attention to membership awareness and collaborative preference specification. However, the column "decision awareness" is implemented in only three systems. Thus, the research in group decision making is still insufficient. This motivates us to explore novel social interfaces for group recommenders. We present our design solutions and user studies in Chapter 3.

2.3 Interaction Design for Group Recommender Systems

Support for such collaboration has been mainly provided by tools that visualize membership awareness, preference awareness and decision awareness. However, these mechanisms do not address group dynamic issues: how member may affect each other. Besides supporting collaboration, members' social relationship and interaction dynamics [69] are also identified as important dimensions for group recommender systems. For example, recent study results show that the more a user is familiar with other group members and the more s/he trusts them, the more s/he will like the items they proposed [106]. Other group members' judgments and emotions can also influence an individual's satisfaction in a group recommender [90] and thus presenting other members' emotional states towards recommended items would make it easier for users to accept items they do not like. However, researchers have not paid attention to the active design of affective interfaces that may influence group dynamics and improve group satisfaction. We investigate these issues and report the findings in Chapter 3.

Researchers have identified the leverage of social interaction as an essential element in designing persuasive technology [29, 41, 113]. Social interaction, such as peer-support, competition, cooperation, has been a clear motivator for wellness activities [20, 85]. Sharing physical activity data with buddies who exercise together have been reported to be effective in motivating users. This is reflected in users' qualitative feedback from the study of both Consolvo et al. [41] and Toscos et al. [113]. Currently, many mobile apps allow users to share their physical activities using social media. Typical examples include Nike+ [8], RunKeeper [11], Runtastic [12], etc. Munson and Consolvo [96] showed users' hesitation in sharing in social networks due to the concerns about whom to share and over-sharing.

Besides the above social interaction schemes, researchers also uncovered the motivating effect of social communication. Such benefits include 1) providing a channel to give support to group members [41], 2) raising the awareness of each other's effort [95], 3) increasing a user's responsibility to exercise [22] and 4) enhancing players' intimacy, which builds a solid basis for long-term play experience [29]. But the above findings were mainly collected from qualitative user feedback. Despite the evidence that shows the promising effect of designing social interaction in persuasive technology, little is found on how to design such social interactions and how one interaction scheme differs from another. This motivates us to conducting in-depth studies, which is reported in Chapter 4.

2.4 Social Interface Design

Conventional social interfaces are mainly concerned with capturing and delivering information among group members. However, less attention is paid to how the group members dynamically interact during collaboration and how their social relationship may dynamically evolve with during the interaction. To this end, we surveyed the literature and presented five most typical systems that aim to fill the gap in traditional interfaces: ***Meeting Mediator***, ***Chalk Sound***, ***CoupleVIBE***, ***CoDine***, and ***Social Camera***. We present the system features, report findings from user studies, and derive corresponding design guidelines.

Meeting Mediator Meeting Mediator (MM) [79] is a mobile system – and one of the few systems – that aims at enhancing group collaboration by detecting and visualizing social interaction and dynamics within group members in real time. More specifically, MM addresses the following problems in group collaboration: alleviating social loafing, reducing production blocking and lessening incomplete information exchange. It recognizes social interaction using a sensing component called Sociometric badge that extracts speech features, measures body movement, detects proximity and captures face-to-face interaction. Based on the above signals, sociometric badge analyzes social interaction and communicates with MM on the mobile phone, which visualizes the social interactions and dynamics in the peripheral during the group work. It uses squares on the four corners of the mobile interface to represent each participant and a circle in the middle shifts dynamically towards the more active member. When the group discussion is balanced among the four members, the circle remains green; otherwise, it turns white.

An in-laboratory controlled user study with 36 groups (each with 4 group members) was conducted to evaluate the effectiveness of MM in both collocated and distributed collaboration groups. The study consisted of two sessions: a problem-solving session and a brainstorming session. Effectiveness was elaborated with two major criteria: *group performance* and *cooperation coherency*. Their results showed that groups who used MM were more efficient in problem-solving session and generated more ideas in brainstorming session. Moreover, MM effectively detected the dominance among the group, which is associated with duration, length, energy and body movement during the speech. MM thus successfully balanced the activeness among the group members. Visualization of group dynamics in the peripheral display is found to be an effective means of persuasion. Their work indicates that the intervention on group dynamics can improve the satisfaction and group performance in collaborative group work. Based on the findings of Meeting Mediator, we summarize the following design guideline:

Design Principle 2.4.1. *Consider providing sufficient information to enhance awareness of group dynamics during collaboration.*

Chalk Sound Awareness of group members' activities is crucial for collaboration. However, mere visual presentations are limited in the "seeing" level of awareness. Chalk Sound [63] integrates dynamic audio as a complement for visual view for off-screen activity in a groupware workspace. It uses speed and pressure of a stylus input on tablet and synthetic audio as output. The audio was captured from real sound.

The system was validated using two studies. Study 1 compared synthetic audio and real sound with 14 participants in an laboratory setting. The researchers synthesized audio sound for drawing various shapes and let participants guess the corresponding shapes. The results showed that real-world sound was more accurate than synthesized sound. Study 2 evaluated the effectiveness of audio awareness with 12 participants in a real collaboration environment. The study compared the visual-only and the visual-with-synthesized-audio conditions. Results showed that 10 out of 12 participants preferred the condition of visual with synthesized audio. These studies demonstrate the feasibility and merits of extending awareness beyond visual modality. Thus, we provide the following design guideline:

Design Principle 2.4.2. *Consider providing multiple modalities of information to enhance the level of awareness among group members during collaboration.*

Traditional social communication systems mainly deal with exchanging *information* among communication parties. Social interfaces can also be designed to help users form, retain and enhance social relationship during communication. This type of communication is also known as *phatic interaction* [117]. The following systems demonstrate how to design social interfaces that go beyond *awareness* and address *connectedness* [52], defined as “a positive emotional appraisal characterized by a feeling of staying in touch within ongoing social relationships." [72]

CoDine CoDine [124] is an intelligent dining table that integrates multi-sensory systems to augment and deliver dining experience for distributed families. This system targets at creating co-presence and enhancing connectedness among family members instead of fulfilling a concrete function or improving efficiency. The system chooses dining activity, which is considered as an indispensable aspect of daily life as well as an important occasion for family gathering. CoDine uses physical object because tangible media can bring extended channel to enhance emotional connectedness. CoDine consists of an interaction screen that supports gestural input, a hosting table that holds food, an ambient table cloth that displays affective messages and a food teleportation device that prints messages on the food.

The system is tested through an in-lab study with 29 participants [123]. Users’ subjective feedback indicates the following potentiality of CoDine system: 1) providing the feeling of co-dining; 2) augmenting the sense of being-together with tangible interactions; 3) increasing the social bonds for non-collocated families and 4) enhancing engagement between co-diners. Based on the design and user study findings, we provide the following design suggestion:

Design Principle 2.4.3. *Consider providing awareness for the common and frequent activities to establish the feeling of connectedness.*

Design Principle 2.4.4. *Consider providing multi-modal sensory information to enable a wider range of emotional interaction.*

Design Principle 2.4.5. *Consider physicality to increase the feeling of connectedness and shorten perceived physical distance in distributed online group interaction.*

CoupleVIBE CoupleVIBE [24] is a mobile application that targets at increasing awareness between couples using their location information. It automatically pushes a user’s location information to the partner’s mobile phone via vibrotactile cues. It aims to promote the connectedness and intimacy by providing couples a way to “think about each other,” i.e. where they are and what they are doing.

The system design follows three principles [24]: 1) balancing between meaningfulness and user control – automatically learning the useful location information to avoid over-sharing while enabling users to control location sharing to protect privacy; 2) balancing between unobtrusiveness and reliable delivery – using vibrotactile cues so that it will less distract users but not too obtrusive to guarantee that the receiver can get the messages; 3) making design decisions consistent with the way physical implicit communication works.

A field study was conducted to examine the roles of CoupleVIBE in enhancing intimacy. Seven couples used the application for 2 – 3 weeks. Users reported CoupleVIBE as useful tool to coordinate their work by showing their availabilities, increase their sense of connectedness, and assure each other by knowing the partner is safe and sound. Similar with CoDine, CoupleVIBE shares *seemingly insignificant information* between the participants, i.e., co-diners or couples, to increase users' knowledge about group members and hence increases their intimacy. It aligns with the CoDine and ChalkSound which use information channel beyond graphical user interface. Thus, we suggest:

Design Principle 2.4.6. *Consider providing awareness for the "seemingly insignificant information" to enhance users' interpersonal relationship and perceived connectedness.*

Additionally, the system thrives to make sense of sensory information and balance between information sharing and privacy, and balance between simplicity and understandability. Thus we summarize its design guidelines below:

Design Principle 2.4.7. *When designing information-sharing features, consider balancing between accuracy and user-control, and between unobtrusive sensing and reliable deliveries.*

Social Camera Social Camera [45] is a mobile social-photography service that allows users to provide implicit feedback using emotional response for photos among non-collocated close-knit groups. This work is another example of creating awareness for phatic communication and consequently increasing intimacy. The motivation of the system is to re-affirm relationships rather than sharing information. When users browse photos posted by their friends in Social Camera, the system automatically captures users' facial expression, which is considered as a spontaneous emotional response, and shares in close-knit groups. Once Social Camera has recorded a video, users have the chance to accept the video, reject it, trim it, make a new one, or not to share it.

Four groups of users, each consisting of four to six members, have tested Social Camera during a field study for two to four weeks. Users shared nearly half of the photos they had taken and they usually provided comments right after seeing a posted photo. In the post-study interview, users widely acknowledged the value of sharing emotional responses in supporting the sense of co-presence for non-collocated close-knit groups. They felt that providing feedback on the photos connected the taker of a photo and others. The users also expressed the merits of personalized expression using video-recorded facial expressions compared with "Like" button, which makes everyone's feedback the same.

This system conforms with the design principles about supporting user control of shared information and content. It also shows how *seemingly trivial information*, such as emotional feedback, helps users to stay connected.

Design Principle 2.4.8. *Consider using emotional feedback to enhance the sense of co-presence and intimacy among non-collocated group members.*

Additionally, since such information is *seemingly insignificant*, it is essential that the input of the information is designed in a light way. In other words, the design principle can be summarized below:

Design Principle 2.4.9. *When providing awareness feature, consider using lightweight input method to avoid excessive user effort, which may lead to under-contribution.*

Summary This subsection presents an in-depth study of five groupware systems for the group collaboration and communication. Compared with traditional social interfaces for groupware, these systems all aim at making the *richness* and *process* of computer-mediated social interaction closer to face-to-face interaction.

In collaboration, we address providing **richness** of interaction channels to reduce the gap between online and offline interactions. We also emphasize the benefits of taking into account **group dynamics** to create a healthier online collaboration environment.

As for communication, the take-home message is to go beyond sharing information, and creating interfaces that aim to enhance **connectedness** and **intimacy** among group members. One common practice is to enable users to share seemingly trivial yet indispensable daily activities with one another. The information should be deployed with sufficient richness, provide input in a light way, and support user control.

2.5 Social Interaction Design

In this section, we present five typical systems that employ the social persuasion theory in designing technologies for behavior change. We characterize the systems into two levels of social support: social awareness and social influence. Social awareness refers to being aware of each other's progress in the behavior-change process, while social influence refers to *explicitly* integrating social interaction mechanisms in the system to motivate healthy behavior change. We show five most typical systems: ***Playful Bottle*** for forming healthy water-drinking, ***BuddyClock*** for sleep management, ***Fish'n'Steps*** and ***Jogging Over Distance*** for physical activities, and ***VERA*** for making health choices.

BuddyClock BuddyClock [78] is an alarm clock on a mobile tablet that automatically shares users' information about the sleeping behavior within a small social network. A user's sleeping status is labeled as "asleep", "awake" or "snoozing". Users can communicate with each other via a messaging component. Notifications are delivered through ambient sound. BuddyClock aims at promoting healthier sleeping habit by providing awareness and enhancing intimacy among shared members.

A study was conducted among five groups of users who used BuddyClock for three to six weeks. The study revealed three interesting findings of BuddyClock. First, it effectively enhanced intimacy among group members – they felt being closer, and got to know the sleeping patterns of each other. Furthermore, being familiar with each other's patterns led to group conformity in sleeping behavior – users became similar in sleeping and getting-up time. Third, users were influenced by each other's sleeping behavior when making choices. Some of them called their buddies when they detected they were still asleep in the morning; others became better at deciding whether to call each other in the evening. Awareness of sleeping time also helped users to reflect on their behavior. The study showed that users were not concerned much about their privacy for sleeping information among close-knit groups.

The study confirms the Design Guideline 2.4.6 that sharing the seemingly trivial information can increase the intimacy among group members. Furthermore, sharing each other's sleeping status can create invisible conformity that helps users to form similar sleeping patterns. This helps us to derive the following design guideline:

Design Principle 2.5.1. *Consider designing awareness for persuasive technologies so that users can implicitly influence each other.*

Jogging Over A Distance Jogging Over a Distance (JOD) [95] is a social jogging system designed for non-collocated dyads to create co-jogging experience. It uses mobile phones and headsets to support conversations. Different from other systems for behavior change, JOD emphasizes on temporal awareness and conversation between running buddies during the jogging process rather than providing post-exercise feedback [95].

JOD consists of three components: 1) a heart-rate sensor for collecting physiological data as input for social comparison, 2) a mobile phone attached to a waist belt communicating with heart-rate sensor in real-time, and 3) a headset that delivers audio feedback about running status of the jogging buddy. Instead of measuring their absolute speeds, JOD compares their heart-rates, maps them to relative distance, i.e. who is in front and who is behind, and translates the relation to spatial audio sound. In this way, if a user is paying greater effort in the jogging – with higher heart rate – the audio mimics the sound as if the user is ahead of the running buddy.

Seventeen users participated in a total of fourteen runs in a field trial. Results of post-study interview led to three major suggestions for designing social experience for jogging applications: 1) providing appropriate effort comprehension mechanism between exercising dyads, e.g., their exertion investment; 2) translating exercise effort to digital representation during the jogging; 3) integrating communication components to link users' physical effort to communication channel.

Overall, their study showed that social presence and social influence serve as a motivator for exertion buddies to jog together despite geographical separation. This confirms the Design Guideline 2.4.3 that embedding multiple channels of awareness can create the sense of co-presence, even if in the process of jogging. Another prominent contribution is the design of its communication component that uses temporal information to enhance the sense of connectedness. Thus, we summarize the following design principle:

Design Principle 2.5.2. *Consider integrating communication component for users to actively persuade and influence each other; when designing communication components, consider designing temporal cues to enhance users' perception of connectedness.*

VERA Virtual Environments for Raised Awareness (VERA) [25] is a mobile application that allows users to make healthy decisions and share in the social context. Users can take photos about an action related with health decision, claim whether they did this or not (accomplishment), append comments, provide rating on how healthy the action is (self-evaluation), and share them in the social network. VERA not only aims at raising users' self-awareness about health decisions, but most importantly, also sustaining user engagement in a social environment.

Two studies were conducted to evaluate VERA. Study 1 compared individual and social usage of VERA. In social usage, participants used the application with other users in the study (mostly strangers) simultaneously. Qualitative results showed that users gained more knowledge about healthy behavior from other people in social usage; meanwhile, users concerned with their privacy when using it among strangers and preferred to use with familiar people. Study 2 compared the effectiveness in groups with shared goals and those without. Results showed that number of statuses and messages of users in both conditions reduced from the first week to last week. The difference between the two weeks was significant in groups without shared goals but not significant in groups with shared goals. Two types of interaction among group members emerged from the study: competing with and accountable for others, despite anonymous usage. Shared goal also increased the feeling of group-belonging, which could be explained by the phenomenon that people were desired to be in a group of people "more like me".

The study reveals a few design suggestions. First, social awareness helps users to sustain in health behavior. Second, system should afford varied degree of sharing and protect user privacy. Third, system should help users find suitable buddy to share. For instance, people trust users if they are similar, e.g., they know each other, or they have common goals. Last but not least, systems should support social interaction between group members, or individual to group relations as they termed. Thus, we propose the following design guidelines:

Design Principle 2.5.3. *Consider providing awareness for users' activities in a social community to enhance user engagement in the long run.*

Design Principle 2.5.4. *Consider helping users find buddies with similar health goals to facilitate group persuasion.*

Design Principle 2.5.5. *Consider affording different types of group interactions, such as interactions among group members and group-belonging.*

Playful Bottle Playful Bottle is a mobile system that uses a mobile phone attaching to an everyday drinking mug to help office workers form a healthy habit of everyday water-intake [39]. Drinking behavior is detected by sensing the motion of the mobile phone using its built-in accelerometers. The quantity of water intake is detected by scanning and comparing the pattern bar on the drinking mug. The Playful Bottle is then integrated into two games: *Tree Game* and *Forest Game*, which allow users to monitor their drinking water in individual game mode and group game mode respectively. In both games, the system automatically sends reminders to users, which is referred to as *system reminder*. In Forest Game, users can persuade to each other, which is called *social reminder*.

The Tree Game and Forest Game are evaluated through a field study (N = 16), including a three-week pretest and a four-week intervention. The goal of the evaluation is to compare the intervention effects of three settings in motivating behavior change: monitoring water intake using sensors and system persuasion (Playful Bottle with Tree Game), and monitoring and social persuasion (Playful Bottle with Forest Game). The study evaluated intervention effectiveness using two metrics: water intake volume and frequency of drinking. The study further compares the persuasion effect from system and peers using response time in drinking water after receiving reminders. Results showed that Forest Game outperformed Tree Game, while both of which were superior to the controlled version, i.e. monitoring only. Thus, we propose the following design principle:

Design Principle 2.5.6. *Consider designing social games to increase intervention effect in behavior change.*

Besides intervention effect, the study results also revealed patterns in dynamics and conformity in a social environment (i.e. Forest Game). During the course of the game, participants in a group tended to evolve to the roles of either care-givers or care-receivers. Particularly, those who acted as care-givers tended to have higher motivation in improving their own water intake. Conformity is another interesting discovery by examining the standard deviation of water intake volume among group members in control phase and intervention phase. They found increased congruence in the amount of water-intake in the group from the control to intervention phase, which could be explained by the theory of social conformity in behavior change. According to their findings, we summarize the following design principles for social interfaces for behavior change:

Design Principle 2.5.7. *Consider integrating social persuasion for behavior change. More concretely, consider designing care-giving, gift-giving and persuasion for social persuasion.*

Design Principle 2.5.8. *Consider enforcing positive group conformity within groups to enhance the intervention effect of behavior change.*

Fish'n'Steps Fish'n'Steps [85] is a social computer game that combines a pedometer to motivate users to do physical exercises. Fish'n'Steps links a player's daily step count to the growth and activity of fish in a fish tank. The fish will show a happy face when the user reaches sufficient progress, an angry face for almost sufficient progress and a sad face for insufficient progress. Users can play alone or with buddies. When playing with multiple users, they can see their fish living in the same tank with other fish, which represent the performance of other players. While rewarding users with the growth of the fish, the game also provides punishment, such as dirty water or removal of decoration if a user fails to meet the sufficient physical activity levels. Therefore, visualizing performance of activities in a fish tank helps users compare and compete, and maintaining a beautiful fish tank allows users to cooperate with each other.

During a 14-week long field study, 19 users participated in three phases: four weeks of pre-study using a pedometer, six weeks of intervention using Fish'n'Step, and four weeks of post-study using the pedometer only. The study aims at evaluating the effectiveness of the application in enhancing users' physical activities and changing their attitudes about exercising. Results showed users increased their steps by using Fish'n'Steps, indicating the effectiveness of the intervention. However, the study does not show a significant difference between individual mode and group mode.

Users have mixed attitudes towards competitive aspects of the game. Some felt competition an important motivation; some felt it incompatible with the spirit of the game. Other users simply did not like competition due to the excessive competition in life. However, cooperation did not significantly improve users' performance because the group members were anonymous. Most participants felt awkward about contacting anonymous users. Different from the results from Playful Bottle [39], the study did not show any significant difference between individual game and social game. But the study demonstrated that game elements, such as using the metaphor of fish and fish tank, could effectively motivate users to be more physical active. Thus, we provide the following design guideline:

Design Principle 2.5.9. *Consider designing gamification elements to enhance user engagement.*

2.6 Design Guidelines

User experience has been a buzz word today, and social interfaces go beyond user experience and address ***co-experience***, as proposed by Forlizzi and Katja [57]. Co-experience is defined as user experience in social contexts when experiences are created together or shared with others. Social interfaces may enable co-experience by providing new channels for social interactions. On the other hand, poorly designed interfaces may be "socially ignorant" [102]. Pentland [102] is one of the earliest researchers aiming at building groupware that detects and visualizes social context during online communication. He addresses the importance of enriching communication context by quantifying social signals such as users' tone of voice, facial movement or gestures. This could be further used to improve group communication and provide real-time intervention. Tran et al. identified ***awareness*** as a crucial element to support natural and effective communication [114]. They further proposed four aspects of awareness in instant messaging systems: presence awareness, turn-taking awareness, contextual awareness, and emotion awareness.

Dey et al. argue that social systems should go beyond awareness to ***connectedness*** [52], defined as “a positive emotional appraisal which is characterized by a feeling of staying in touch within ongoing social relationships." [72] They further identify connectedness as a social affordance in computer-mediated communication, but unfortunately largely neglected. They proposed four design guidelines for peripheral displays to enhance connectedness, including designing interfaces that are physical, peripheral, small and choosing the social-bonding artifacts.

Hassenzahl et al. [66] survey techniques that aim to create the experience of ***relatedness*** – love, closeness and intimacy, and present six strategies: awareness, expressivity, physicalness, gift-giving, joint-action, and memory. Among these strategies, they consider awareness as the most comprehensive strategy to create a feeling of cognitive awareness by sharing different types of information, including presence, activities, or mood among a group of people. This conforms with the design principles we derived earlier in this section (Design Guideline 2.4.3). A related strategy is expressivity that enables partners to express their feelings and emotions (as is reflected in Design Principle 2.4.4). Physicalness aims to complement the missing dimension of physical touch in online communication (Design Principle 2.4.5). However, physicalness is identified as one of the most challenging strategies among the six. Gift-giving strategy demonstrates users’ caring and valuing the other person through gift-giving. This confirms Design Principle 2.5.7. Joint action allows users to carry out an action together. Intuitively, this strategy requires the users to be physically collocated. Many systems have been designed to support co-actions in non-collocated situations. It requires appropriate communication components, behavioral interdependence and proper activities (see Design Principle 2.5.2). When designing this strategy, it is essential to select the activities that retain interdependence. Hassenzahl et al. consider it desirable to choose everyday, mundane, and real-world activities which are more natural. Finally, memories enhance the feeling of relatedness by keeping records of past, present and future activities.

The above work mainly covers three requirements for social interface design: *awareness*, *connectedness*, and *relatedness*. In this thesis, we aim to present more comprehensive perspectives and guidelines for social experience design.

3 Social Interface Design for Group Recommenders

In this chapter, we study how to design emotion-awareness tools as social interfaces for group recommender systems. Part of this chapter is published in [34, 35, 36, 33].

3.1 Introduction

In Chapter 2, we survey the current literature about interface and interaction design for group recommender systems. We identify mutual awareness among group members as an essential technique in social interface design for group recommender systems. More concretely, current group recommenders mainly apply three types of mutual awareness: membership awareness, preference awareness, and decision awareness. However, the emphasis on emotion awareness among group members is lacking. To study the roles of emotion awareness, we implemented a group recommender system called GroupFun. We further designed three different emotion awareness tools in GroupFun. Study results showed that emotion awareness tools have the potential to help users know better about each other's preferences, enhance their social relationship, increase the feeling of connectedness, and help them to influence each other to facilitate decision-making. Finally, we derived design implications for designing social interfaces – especially emotion-aware interfaces – for group recommender systems.

3.2 Related Work

3.2.1 Emotion Awareness Tools in Computer Mediated Communication

Emotion awareness is not a new concept in groupware systems. A typical application area is computer mediated communication (CMC). Emoticons are an important kind of language of emotions utilized in CMC. They are approximate facial expressions simulated by a combination of certain keystrokes, and they are widely used in online communications, e.g. instant messaging and emails [118]. Studies [107] show that users who used emoticons in online communication were more satisfied than those who had no access to emoticons. Concretely, emoticons help users to perceive the correct emotion, attitude and attention intents of received messages [86]. There has been some work on animation to visualize emotions. Wang et al. [119] prototyped a chatting system that animated text associated with emotional information to show users' affective states. Their study results show that emotional sharing improves the accuracy of the perception of each other's meanings and enhances the user's chatting experience. Another dimension for communicating emotion is haptic feedback. A typical example is EmoHeart [97], which interprets user emotions through chatting messages and transmits them through an online avatar coupled to a haptic simulation apparatus. Cui et al. [45] used video recordings to deliver emotional responses and found emotional responses can enhance social interaction in close-knit groups.

However, in a group music recommender system, users listen to music continuously over 3–4 minutes, while in CMC users exchange information in a much shorter time frame. The different requirements for users' attention and experience flow drive us to explore other possibilities for communicating emotions in group music recommenders.

3.2.2 Emotion Categories

There have been various ways to represent emotions. The best-known model is Ekman's [53], represented by 6 basic human emotions (anger, disgust, fear, happiness, sadness, and surprise) and their relations to facial expressions. The Geneva Emotion Wheel [109] evaluates emotional responses related to objects, events and situations; it is a more refined model of emotion, incorporating 20 emotion categories. Besides these two general models, researchers have provided domain-specific emotion models, such as emotions related to products [50], visual interfaces [71], music [129], films [61] and pictures [27]. The above research shows that domain specific emotions can provide more accurate descriptions of users' emotions in the corresponding contexts. Most of these emotion categories are visually represented by facial expressions and body language [49, 71] using anonymous and non-personalized figures.

3.3 GroupFun – a Music Group Recommender System

To investigate social interface design, we developed a music group recommender system called GroupFun (see Figure 3.1). We first conducted a need assessment survey for music group recommender systems and derived user goals and requirements for GroupFun. We then designed and developed GroupFun based on results of user research.

3.3.1 Need Assessment Survey

We performed need assessment interviews with 11 users who use social network services on a daily basis. These users are from six different countries; seven of them are university students. When asked how they discovered music, eight of them mentioned that they would like their friends to recommend music. Nine of them are willing to share music playlists with their friends, ten are interested in knowing friends' music taste and ten would like to mutually share music with their friends. When asked about how they would like to share music, all of them expect some applications to help them automatically fulfill this task. In the meantime, four of them also pointed out the benefit of face-to-face discussions and three also preferred online chatting tools to assist music sharing. This indicates users' need for communication in GroupFun. When we asked them what would be the most attractive features for such applications, the following factors were emphasized: 1) attractive interfaces, 2) application enjoyability, 3) ease to learn and ease to use, 4) saving effort, and 5) discovering new music.

3.3.2 Interface and Interaction Design

Based on the interview results, we have developed a music group recommender system named GroupFun, which is a Facebook application that allows groups of users to share music events, such as a graduation party. The functions of GroupFun mainly include group management, music uploads, music rating and group music recommendation. Users can log in to GroupFun with Facebook accounts, create and join a group, invite members, upload music, and listen to a common playlist. They can specify music preferences by rating the songs provided by GroupFun. GroupFun will generate music playlists by aggregating group members' ratings. Figure 3.1 and 3.2 show the home page and group management page respectively.

We followed two principles when designing group interface: 1) enhance mutual awareness and 2) enhance transparency. Screenshot of the group song page is shown in Figure 3.3. The group song page shows the list of recommended group songs ranked by the recommendation values. This page achieves membership awareness by listing all members and displaying their Facebook profile pictures in the group. Preference awareness is implemented by enabling members to rate the group songs. Since listening to music requires less effort in decision-making, compared with traveling, movies, etc, we provide a light-weight decision-awareness feature – by enabling users to check who have rated certain music and the rating values.

GroupFun generates recommendations by aggregating all users' preference using Probabilistic Weighted Sum (PWS) to maximize user satisfaction and reduce the chance of manipulation [104]. We do not extensively discuss the algorithms since it is out of scope of this thesis.

Based on users' requests on sharing music on multiple devices, we developed a corresponding mobile version of GroupFun on Android platform. Users can create groups, invite Facebook friends to the groups, rate music to indicate their preferences, and listen to group songs. The interfaces is shown in Figure 3.4.

Figure 3.1: GroupFun interface: Home page.

3.3.3 Summary

We employ GroupFun as an experimental platform to study social interfaces for group recommender systems. It implements mutual awareness for membership, preference and decision-making. In the following sections, we introduce emotion-awareness features and present the design, implementation, and evaluation of such interfaces in GroupFun.

3.4 ACTI – Affective Color Tagging Interface

In this section, we designed an emotional wheel, which enables users to tag their emotional attitudes after listening to each song. We aim to design an emotion awareness tool that meets

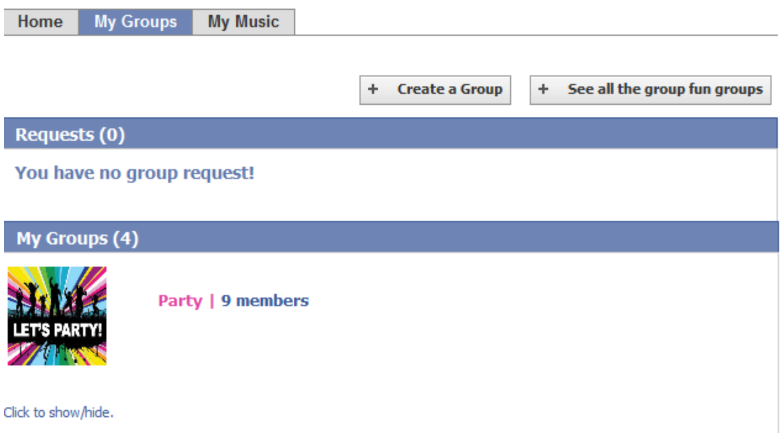


Figure 3.2: GroupFun interface: My Groups.

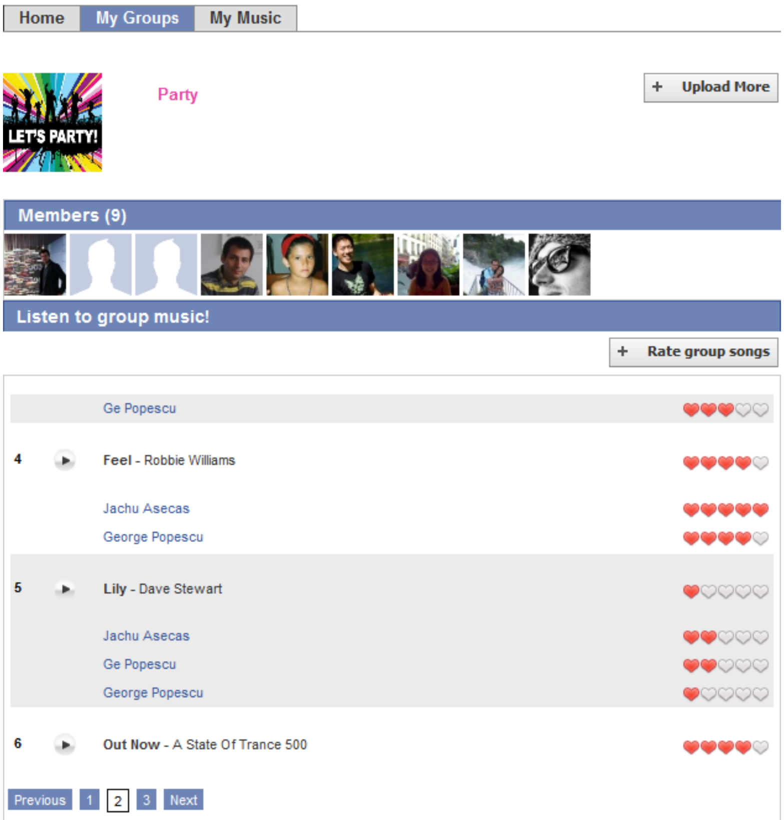


Figure 3.3: GroupFun interface: Group songs.

the following requirements:

1. **Exploratory and collaborative**: allowing users to discover the preferences of other participants and construct suitable preferences together;.
2. **Flexible and incremental**: accommodating users' individual preference structures and allowing reasonable changes in decisions to be made and incorporated in the current problem solving process.

3.4.1 Interface Design

In this section, we focus on emotions related to music. We selected the nine emotion categories from the Geneva Emotional Music Scale (GEMS-9) [129], which is considered as the first tool devised to measure music-evoked emotions. We adopt the short version GEMS-9, consisting of nine classes of emotions, including *wonder, transcendence, power, tenderness, nostalgia, peacefulness, joyful, sadness, tension*. Each class provides a scale from 1 to 5 indicating the intensity of the emotion. Table 3.1 explains these nine categories.

However, asking users to evaluate the evoked emotion of music is not our research focus.

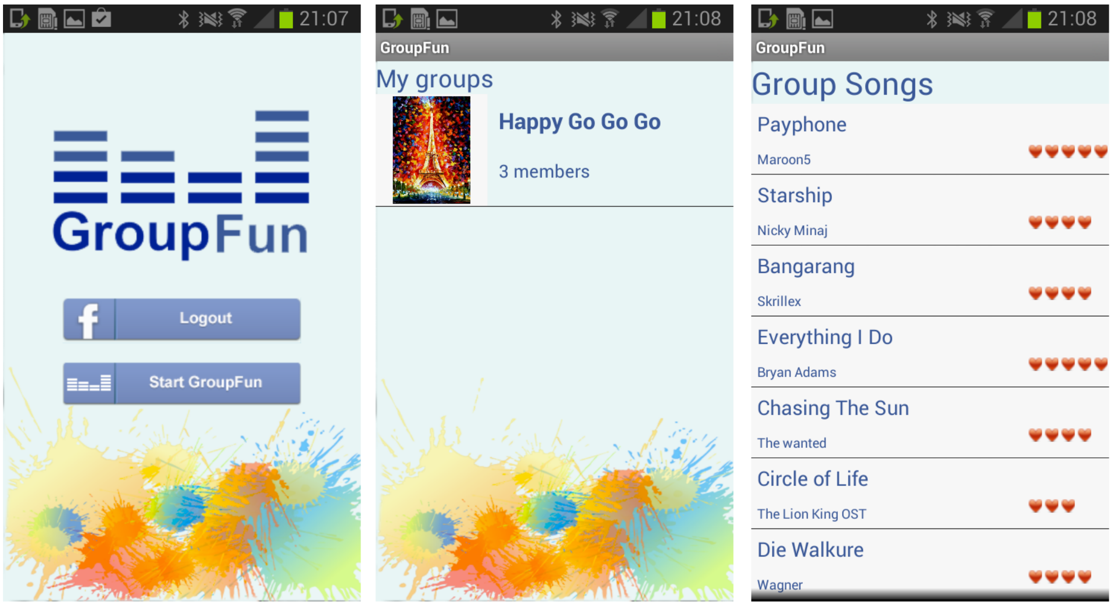


Figure 3.4: GroupFun interface: mobile application.

Rather, we want to provide a ludic affective feedback interface for a group of users to participate. A survey-style questionnaire easily distracts users from the system itself. Inspired by Geneva Emotional Wheel (GEW) [109] (Figure 3.5), we visualize the evaluation scale to a color wheel. Figure 3.6 shows the graphical design of ACTI.

ACTI design incorporates emotion categories of GEMS as 'body,' Geneva Emotion Wheel as 'skeleton' and 'clothes.' The outcome of the design is shown in Figure 3.6 . There are nine emotions in the ACTI, corresponding to the nine emotion categories in GEMS, namely joyful, nostalgia, tenderness, peaceful, power, transcendence, wonder, sadness, and tension. Each emotion contains five levels of intensity, inspired by Geneva Emotion Wheel. The levels of intensity are distinguished by the sizes and color transparency of the scales. Please note that the colors are not designed to represent specific meanings, but to make the interface wheel look fun and appealing.

ACTI meets the criteria we proposed: exploratory and collaborative, flexible and incremental. First, ACTI allows collaborative tagging and supports group emotion visualization. Second, as music contains high degree of emotional elements [18], it is more convenient to comment a song by emotional tagging. Third, one song might suit for a limited number of themes, but the suitability is difficult to characterize by ratings; instead, the emotions contained in a song are more stable, which could be used to assess contextual suitability. The combination of emotion and color tagging leads to an interface with inspiring visualization and ludic feedback. We will evaluate ACTI by comparative user studies.

We further include ACTI widget in GroupFun rating interface. Users could give emotional feedback to songs by clicking the emotional button at the left side of song ratings. The emo-

Category	Explanations
Wonder	Happy, amazed, dazzled, allured, moved
Transcendence	Inspired, feeling of spirituality
Tenderness	In love, affectionate, sensual, tender
Nostalgia	Sentimental, dreamy, melancholic
Peacefulness	Calm, relaxed, serene, soothed, meditative
Energy	Energetic, triumphant, fiery, strong, heroic
Joyful activation	Stimulated, joyful, animated, dancing
Tension	Agitated, nervous, tense, impatient
Sadness	Sad, sorrowful

Table 3.1: Emotion Categories in Geneva Emotion Music Scale.

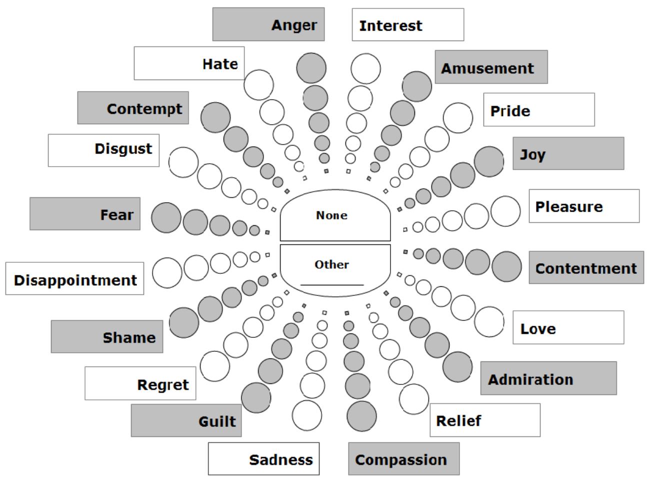


Figure 3.5: Scherer's Geneva Emotion Wheel [109].

tional rating interface will pop out including the collective emotion and individual emotion, as is shown in Figure 3.7.

3.4.2 Pilot Study

Hypotheses

We postulate that interface and interaction design in social group recommender systems depends on the relationship among group members. In other words, interface and interaction design is different in groups constitute of friends and those constitutes of strangers. We also hypothesize that groups whose members are in close relationship would like to use ACTI interface and groups whose members are not close with each other would less likely to.

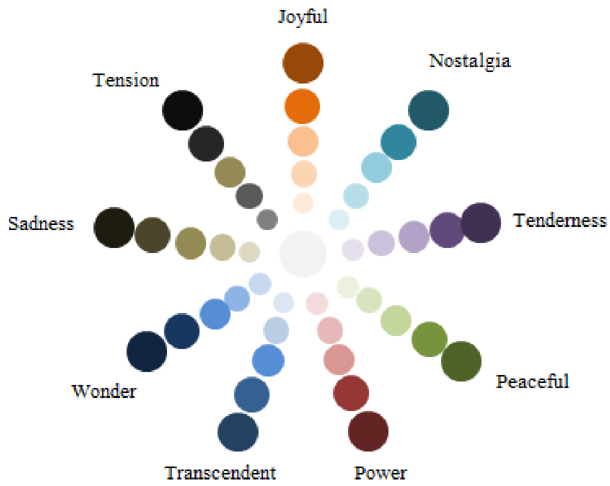


Figure 3.6: ACTI graphical design.

Participants

In total, two groups (four participants each) joined the user study. Two Ph.D. students from a course voluntarily participated in the user study. Each of them invited another three people to form a group. In order to investigate the effect of relationship on group behavior, we asked the first student to invite people who are not familiar with each other, while we asked the second student to invite people who are familiar with each other. We first surveyed their background, including gender, education, age, familiarity with music applications and social network usage. Table 3.2 lists demographic profiles of the studied subjects in the two groups.

	Group 1	Group 2
Gender	Female (1), Male (3)	Male (4)
Education	Master (3), PhD (1)	PhD (4)
Age	20-24 (1), 25-30 (3)	25-30 (4)
Familiarity with music apps	High familiarity	high familiarity
Social network usage/week	5-10 hrs (1), < 5 hours (3)	< 5 hrs (3) 10-5 hrs (1)

Table 3.2: Demographic information of participants.

Scenarios and Roles

We used the car driving scenario as follows. A group of four friends are traveling together from Lausanne to Geneva by car. It is around 45 minutes ride. One of the participants is a driver, and the other three participants are passengers. They are using GroupFun to select a playlist for their trip. This scenario is a typical group activity where all members consume a recommendation list simultaneously. This scenario provides two types of roles, one driver and three passengers. A driver could be the user who creates the group. Since the focus group is

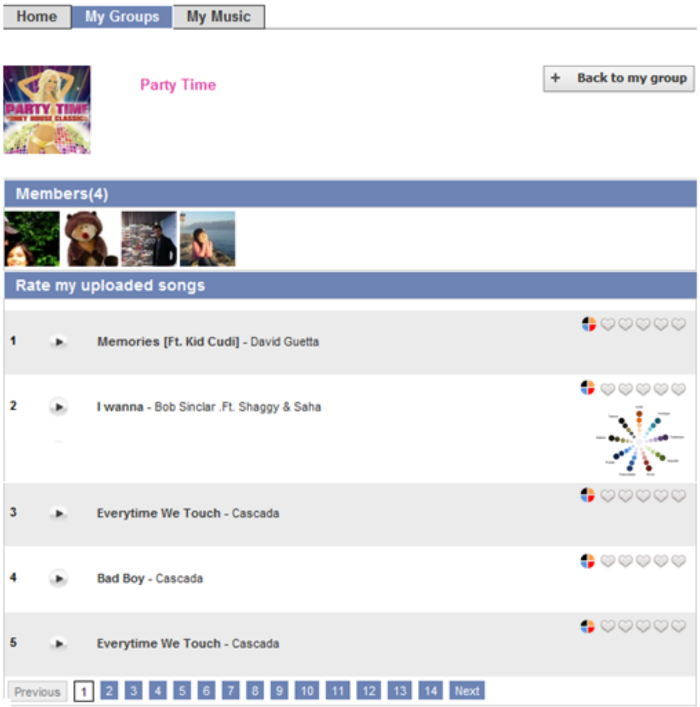


Figure 3.7: Integrating ACTI in GroupFun.

small, it is not realistic to design a large-group scenario such as a party. In this experiment, GroupFun provides 15 songs, which is suitable for the experiment in the sense that 15 songs are approximately suitable for a 45-minute ride.

Experiment Design and Procedure

We carried out the user study on two groups in two different days. We first invited each member of the group to meet in our lab with their favorite music. We debriefed each group on the procedure of the study and the usage of GroupFun. We then assigned roles for each group. Users started following the scenario after exploring original GroupFun for around five minutes. In each group, the driver first created a group and sent invitation to his/her friends (other participants of the group), and uploaded his/her favorite music. Invited users accepted the invitation, joined the group and contributed songs respectively. They could also listen to group music and rate songs. Then they continued with a survey questionnaire evaluating GroupFun. After that, we invited them to explore the experimented GroupFun interface with ACTI, followed by an interview. The user study process in both groups has been video-recorded for further analysis.

3.4.3 Findings

Survey Results

After using GroupFun for the experiment scenario, we first survey users' perception of system quality of GroupFun to assure the quality of recommended songs does not influence users' perception on user issues that we are investigating. We asked users to rate the statements in the survey from 1 (strongly disagree) to 5 (strongly agree). We list the statements and users' average ratings in Table 3.3. From the results, we could see that users were overall satisfied with GroupFun and recommended items. They even indicated their inclination to listen to music recommended by GroupFun again and tell their friends about GroupFun.

Questions	**Mean**
I am satisfied with the songs recommended to me.	3.62
The system suggested an adequate number of songs.	3.75
The playlist is suitable for the whole group.	3.125
GroupFun is useful.	3.875
I will use GroupFun again.	2.87
I will tell my friends about GroupFun.	3.25
I will listen to the group music even when I am alone.	3.75

Table 3.3: Demographic information of participants.

Interview

In order to verify the differences in the relationship among group members, each participant indicated their closeness with each other (except themselves) ranging from 1 (don't know him/her) to 5 (know him/her very well). For example, if Participant A is in a very close relationship with Participant B, he is expected to rate 5 for Participant A. The total score in Group 1 is 33 while that in Group 2 is 44. This result verifies the group differences as we expected.

It is easy to discover from the video recording that both groups enjoyed listening to the group music very much. For example, some of them even sang with the music while listening. More interestingly, we found out members in Group 2 even discussed with each other about a song that they particularly enjoyed.

During the interview phase, we asked them to call back the experiment scenario and whether music influences their mood. All users agreed that listening to the music did evoke different emotions.

Then each participant compared original GroupFun interface with the experiment interface with ACTI. In Group 1, only one member supported ACTI, while others regarded it unnecessary and too complicated to include additional features. It cost them more effort. By contrast, members in Group 2 were highly positive about ACTI. As one user said,

> *"It is interesting to listen to music while evaluating their mood. It is even more interesting to compare their results with group emotion."*

Followed by this question, we asked them whether their emotion would be influenced by others'. As we hypothesized, members in Group 1 did not see any influence, but they still enjoyed using GroupFun because of the recommended group music. By contrast, 3 out of 4 participants in Group 2 said they would like to view group emotion when tagging their own mood evoked by music.

The last question is to ask participants their suggestions on interface and interaction design of GroupFun. We collected some valuable suggestions from participants. One user in Group 1 regarded simplicity as an important factor. He would prefer interesting functions, but the interface should not distract them from the main functions of GroupFun, and that explains why he did not like ACTI. On the other hand, one member from Group 2 said some interesting group activity would help them to listen to suggested songs more carefully and therefore provide more accurate and responsible ratings. Another participant in Group 2 also mentioned that they would like to see more enjoyable interfaces in GroupFun as an entertaining application.

3.4.4 Summary

We designed an affective color tagging interface (ACTI) and applied it to GroupFun, a music group recommender system. We further invited two different types of groups to our user study. While members of one group did not know each other well, the other group consisted of friends in a good relationship. Users in both groups preferred attractive but simple interfaces. Even though both groups of users enjoyed listening to group songs and indicated their positive attitudes towards GroupFun, the user study did show some differences in their group behavior. Members of the group with a close relationship were active in the discussion with each other, and they liked interfaces with group emotion, and considered it interesting and entertaining. Meanwhile, groups whose members were less close with each other considered affective interface less useful. This supports our hypothesis that interface design in social group recommender systems should consider group formation and relationship.

However, this work is still at a preliminary stage and has some limitations. First, as a pilot study, we only invited two groups of users to survey their needs. In order to further establish design guidelines, we need more groups in terms of number and types. Furthermore, web-based affective interface is limited. Rather, user emotion could be captured automatically in an ambient environment. Our future work also include ambient affective interface in social group recommender systems.

3.5 CoFeel – Color Wheels for Emotion Presentation

Based on user feedback from the previous section, we are motivated to design an affective interface that is more engaging and entertaining. Taking advantages of rich sensors on mobile phones, we designed a new version of emotion-awareness tool called CoFeel and integrated it in the mobile version of GroupFun. The rest of this section presents the design, implementation and evaluation of CoFeel.

3.5.1 Interface Design

We aim to design an emotion awareness tool that allows users to accurately tag their emotions and is engaging to motivate users to provide emotional feedback.

Accurate. We choose Russell's emotion complex (Russell 1998)[108] to logically help users annotate emotions into two dimensions: valence and arousal. Valence refers to the polarity of an emotion, i.e., positive or negative, while arousal stands for intensity of the emotion, e.g., *"irritated"* having stronger intensity than *"sad"*. We consider mapping emotions in a dimensional space as a means to quantify the valence and arousal of a given emotion, and thus leading to higher accuracy for annotated emotions.

Engaging. As mentioned before, the emotion annotation tool should be engaging and playful to motivate users to provide emotional feedback. We thus consider Scherer's Geneva Emotion Wheel [109] (Figure 3.5) as a method to visually present emotions in the dimensional space and introduce colors based on Hatt's color wheel [67] (Figure 3.8) to make the interface look appealing.

In essence, CoFeel design incorporates the dimensional presentation of Russel's complex as 'body', Scherer's Geneva Emotion Wheel as 'skeleton,' and Hatt's color wheel as 'clothes.' The outcome of the design is shown in Figure 3.9. After several rounds of design iterations and collecting feedback from users, we propose the CoFeel design with eight emotion spikes, each spike containing five scales and a representative color. We select the eight emotions (*excited*, *happy*, *satisfied*, *relaxed*, *sleepy*, *sad*, *distressed*, and *irritated*) from Russels' complex and consider their positions in the valence and arousal mapping. We are aware that the meanings of colors in representing emotions differ from cultures to countries. We choose the color mapping by considering commonly-acknowledged facts, e.g., yellow for *"joy"* and blue for *"calm"*. We then turn Hatt's color palette and try our best to make the colors match the emotional meaning. Again, the goal of introducing colors is more for visual appealing than for assigning meanings for emotions.

To enhance user engagement in interacting with the CoFeel, we use the metaphor of a plate with each emotion as a hole and a ball rolling on the plate to indicate users' current emotions. Users can change the emotions, i.e., the position of the ball, by rotating and tilting the plate surface (see Figure 3.10). The aim of using the plate-hole-ball metaphor is to enhance user

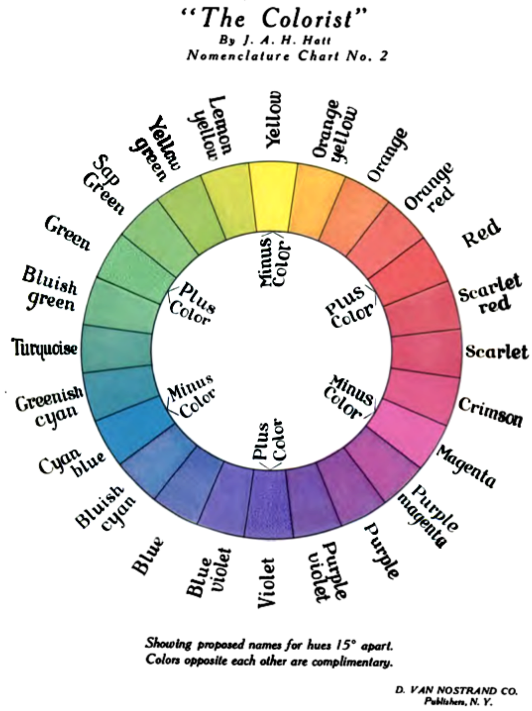


Figure 3.8: Hatt's color wheel [67].

affordance to interact with the interface.

We implement CoFeel emotion plate on Samsung Galaxy SII 9100 with Android OS, but it can be easily used in any other Android devices. The phone detects user movement and direction of surface plate using sensors on mobile phones, i.e., accelerometers and gyroscopes. We have also filtered out constant accelerometer data when users are walking, traveling, etc. In this way, users can input their emotions in a stable way.

We then integrate CoFeel in GroupFun. Figure 3.11 shows the group music interface of GroupFun. The emotion awareness feature consists of two components: individual emotion annotation and group emotion visualization Users can annotate emotions with CoFeel by holding the phone, rolling the ball around the surface of emotion plate, and clicking the "Update Mood" right below the color wheel to confirm annotated emotions. After selecting, the annotation is recorded and visualized in the timeline of the song as group emotions, together with other group members' annotations. The group emotions are visualized as music score (see Figure 3.11). Different emotions are distinguished by colors, corresponding with colors in the color wheel. Intensities of emotions are mapped to the position of the lines from 1 to 5. For example, the beginning of the song 'Fairytale', is rated as an '*relaxed*' song, with the level of 5 out 5.

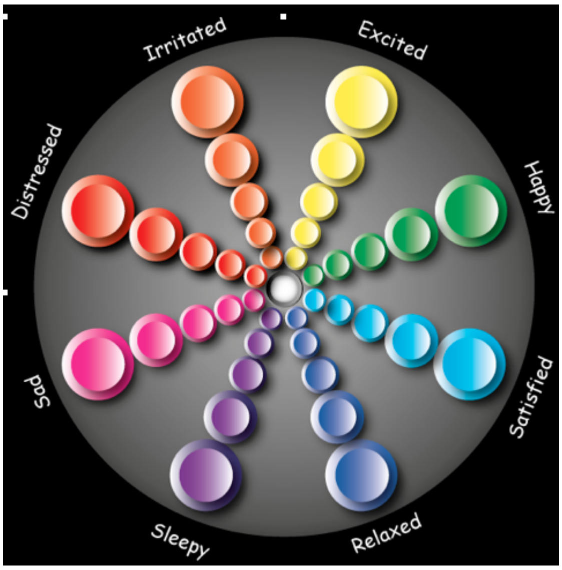


Figure 3.9: Visual Design of CoFeel.

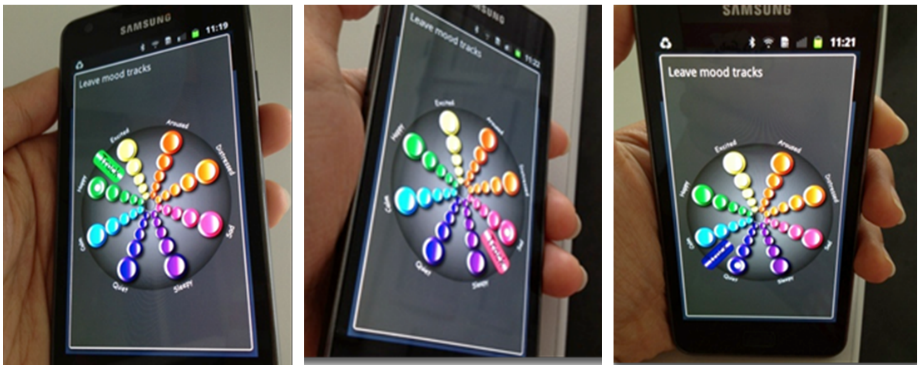


Figure 3.10: Interaction with CoFeel: tilting and rotating.

3.5.2 Pilot Study

To assess the effectiveness of CoFeel and its impact in GroupFun, we conducted a pilot study in our laboratory to make an in-depth observation of the participants and collect their comments when using the system. They were encouraged to think aloud during the study. The process was audio-recorded for further analysis.

Participants and Materials

Four participants, including a student and three researchers, volunteered in the study. All of them had Facebook accounts. They consisted of two females and two males from different countries in Europe and Asia, with their ages ranging from 23 to 29. All of them were currently using smartphones and were familiar with listening to music on smartphones.

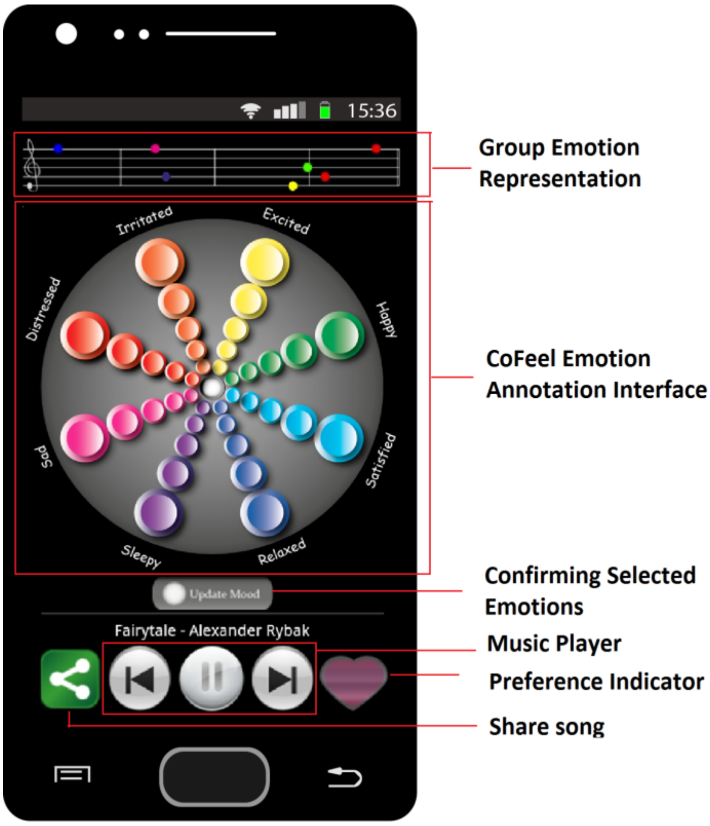


Figure 3.11: Providing emotional feedback to a whole song.

We installed GroupFun on Samsung Galaxy SII 9100 with Android OS 2.3.2. To prepare for the experiment, we selected 20 group songs from top tracks in Last.fm to cover various genres. The length of the songs ranged from 3.5 to 5.2 minutes ($M = 4.2$).

Procedure

Each participant scheduled 30 minutes with us and provided their Facebook homepage via email. We obtained their Facebook ID from Facebook Graph API and created a group in GroupFun for each participant. Each group includes the user and one person randomly chosen from his/her Facebook friend list.

During the study, participants were first debriefed about a scenario in which they (the user and a group member) were recommended with a music playlist by GroupFun and that they could provide emotional feedback that their group members could see. They then logged into GroupFun with their Facebook accounts, entered their group and familiarized themselves with the interface (see Figure 4) using a warm-up song. Afterwards, they started to listen to 5 songs randomly shuffled from the 20 songs that we had chosen. They could annotate emotions during the timeline of the song with CoFeel. This procedure took up to 20 minutes. Finally, they participated in an unstructured interview.

3.5.3 Findings

Overall, users had positive feedback about CoFeel. Meanwhile, they provided further suggestions. We summarize our findings in four aspects as below: accuracy, engagement, emotion awareness interface design, and group influence.

All users agreed that visualizing emotions in a plate of valence and arousal helped them tag emotions. Meanwhile, they felt that the concrete emotion categories could be more dynamic. For example, the word *"irritated"* could also be *"angry"*, *"uneasy"*, etc, depending personal usage. Additionally, the concrete emotional words could vary by context. Thus, it is recommended that CoFeel dynamically update the emotional vocabulary with emotional words that represent equivalent valence and arousal.

Design Principle 3.5.1. *Balance between accuracy and richness of meaning for designing emotion annotation interfaces.*

Users acknowledged the fact that playing with the ball on CoFeel while listening to music was fun. They also pointed out scenarios when the interaction modality might cause inconvenience, such as when moving. Directly clicking on the hole of CoFeel plate could serve as a complementary means to register user emotions.

Design Principle 3.5.2. *Balance between user engagement and convenience for designing emotion annotation interface.*

We also asked for users' opinions on emotion awareness interface, including annotating emotions using color wheel and visualizing group emotions in the metaphor of music notes. Overall, users agree that the interface was understandable, but could be better. A common suggestion is to combine emotional annotation and group emotion visualization. Currently, the two parts are visualized in two different metaphors and arranged separately on the interface. This potentially reduces the perceived relation between individual emotional annotation (input) and group emotion visualization (output). An alternative suggestion is to visualize group emotions dynamically as a layer on top of the emotion color wheel.

Design Principle 3.5.3. *Make the style of individual emotional annotation (input) and group emotion visualization (output) consistent.*

We further asked users whether they expected their emotions to influence that of their peers if their peers had the chance to use the system. Three of them responded positively. Some believed that the number of emotional annotations could reveal their overall preference for the song. Others considered emotions as a means to indicate their attitudes towards the song during the timeline. Meanwhile, one user answered negatively since the group member that we had randomly chosen for her was only an acquaintance, and she would probably not influence him. This confirms with findings in the previous section that group influence is dependent of group relationship.

Design Principle 3.5.4. *Consider the role of social relationship among group members in the effect of group influence.*

We are aware of the small sample size of our user study and that participants used GroupFun individually rather than in a group setting. However, this small-scale pilot study serves two purposes. First, it provided as a first step to investigate and assess designing emotion awareness interface for group recommender systems and provides design implications. Second, it empirically suggests the potential of using emotion awareness interface to enable group influence on preferences in group recommenders.

3.5.4 Summary

We have explored designing emotion awareness features for group recommender systems. We first designed an emotion annotation tool called CoFeel. We then integrated it in GroupFun, a music group recommender systems, which allows users to annotate individual emotions and view group emotions. Results of a qualitative study suggest that CoFeel is a promising tool for emotion annotation and users are likely to influence group members using emotion annotations in group recommenders. This work extends literature in group recommenders by exploring design space for emotion awareness features and their impact in group recommenders. Future work includes continuing with the design and investigating its impact with comparative user studies.

3.6 Empatheticons – Dynamic Pictures for Emotion Presentation

In this section, we describe the design process behind ***empatheticons*** – a set of dynamic emoticons. Moreover, we are interested in understanding how they affect group dynamics and group satisfaction. The following is a motivational example:

> Alicia, Victoria and several of their friends are using GroupFun, a music group recommender, to organize a party tonight. They will rate the songs that the group members have contributed. GroupFun will determine the final playlist. The more actively a member rates the songs, the more her preferences will be taken into account in the final playlist. In the traditional version, members can rate songs using a scale from 1-5 (least to most preferred). In the new version, they can not only provide their ratings but also describe how the music makes them feel using an emotion annotation tool. When Alicia is rating and annotating a song, she can "see" how Victoria and others feel as well as read their ratings.

In the above scenario, we refer to the functionality (which allows group members to annotate and visualize emotions) as ***emotion awareness***. We designed and implemented a set of nine ***empatheticons*** (*for empathy*) for GroupFun based on the kinetic behaviors. Empatheticons represent emotions visually by applying a set of motion schemes to each individual user's profile picture. When multiple participants have annotated the songs with their emotions, GroupFun shows these dynamic emoticons in a continuous and animated way while the music is playing.

Results from two rounds of studies show that users can distinguish and recognize accurately the emotions presented by our empatheticons. In an in-depth laboratory study, participants indicated that empatheticons can enhance user-perceived togetherness and familiarity with other members' preferences, thus increasing their satisfaction with GroupFun. Additionally, users' emotional responses are not only influenced by the other group members, but also serve as a predictor of users' satisfaction for a group recommender.

3.6.1 Interface Design

We were motivated to use the concept of ***kineticon*** [65] – an iconographic motion technique – for visual representation for two reasons. First, research has shown that motion is one of the most prominent types of affective response to music [73], as opposed to text or colors. Second, watching a user (i.e., his profile picture) dance or move brings more intimacy than a neutral avatar. See Section 3 for the design rationale.

Apple Inc. was one of the earliest companies that integrate motion icons. A typical example is an application icon in the dock of the Mac OS. When a user clicks on an icon to launch an application, the icon jumps in the dock to indicate that the application is preparing to launch. Harrison et al. also established a set of 39 kinetic behaviors, such as spin, bounce, running, etc. These kineticons are designed based on the following sources: 1) biological motion, i.e., animal and human body motions; 2) gestures, e.g., head nod, shoulder shrug, thumbs up; 3) organic motion, e.g., blossoming of a flower, beating of a heart; 4) mechanical motion, e.g., toggles, knobs, buttons and sliders; 5) physical and natural effects, e.g., how a leaf falls from a tree or paper folds; and 6) cartoon conventions, which are exaggerated translations of animation to 2D graphical interfaces. These inspirations provided us with guidelines for designing our empatheticons.

Up till now, kineticons were designed to convey the status of a system in a graphical interface. Our work is novel, as it exploits kineticons as a means to visualize emotions.

What is an empatheticon?

An ***empatheticon*** is a set of pre-defined motion frames (or deformations) to visualize an emotion, which could be adapted to any given pictures. Given that every member has a profile picture in GroupFun, we design empatheticons to present each group member's emotions by applying motion features to the profile pictures. Currently, empatheticons present the nine music-evoked emotions in GEMS-9. The empatheticons are implemented as a library on

Android system, with a specified emotion and a picture as input and a kinetic animation as output.

By integrating the empatheticons, GroupFun allows a user to specify his/her emotions when listening to a group song. We herein refer to the process as emotion annotation. The emotions will display as an animated profile picture of the user who annotates. These annotations are temporal and can last as long as the song itself. The user can change his/her emotions and thus a song can have different labels at different time. Group members can see the user's emotions and vice versa. Empatheticons are so far best suited for the following two criteria [49].

- **Intimacy**: an animated profile picture can convey intimacy better than an anonymous and non-personalized emoticon;
- **Immediacy**: 1) they are more vivid because of their animation nature; 2) an animated profile picture can last as long as the song itself; 3) the emotions they express change during the song so they give more emotional nuances.

Choice of Emotion Presentations

We first selected one song for each emotion category as a starting point to design the empatheticons. The music was selected by referring to Zentner's criteria [129] and covered four genres: classical, jazz, pop/rock, and techno. We first identified the mental model of emotion presentation related to music. Besides Harrison's inspiration sources for kineticons, we also referred to the emotion metaphors summarized by Kovecses [81]. The reported emotions include anger, fear, happiness, sadness, love, lust, pride, shame and surprise. We were inspired by the metaphors of anger (for tension), happiness, sadness, love (for tenderness), surprise (for wonder) and pride (for transcendence). For example, we designed joyful activation (abbreviated as joyful) using the metaphor 'leaving the ground and up in the air' and sadness using the metaphor 'lowering down'. For nostalgia and peacefulness, which were not covered by Kovecses, we took videos of people's body expressions in daily life, concerts and parties when they were listening to music. The videos also worked as a source to refine empatheticons that were already designed. Other information sources include online videos, television, etc. Figure 3.12 shows the outcome of empatheticon design, inspirations, descriptions and visual illustrations.

Creation of Empatheticons

We started creating empatheticons by the low-fidelity prototype, i.e., sketches based on deformation to a square. This is followed by high-fidelity (Hi-Fi) prototypes using Adobe Flash. Hi-Fi prototypes also helped us to get feedback from users before implementation. We then implemented the empatheticons by texture mapping in OpenGL ES 10. We chose OpenGL to

Emotions	Inspira-tions	Descriptions	Animation timeline
Wonder	Biological	A person swaying his body from left to right, trying to discover what happened.	
Transcendence	Biological	A light object ascending and descending quietly and slowly.	
Tenderness	Biological	A person rocking the cradle from left to right slowly.	
Nostalgia	Biological	A person breathing deeply.	
Peacefulness	Physical	Water surface changing flexibly and adaptively.	
Energy	Biological	A rugby player performing 'The Haka' dance.	
Joyful	Biological	A person jumping up and down with expectation.	
Tension	Mechanical	An object expanding and shrinking intensely.	
Sadness	Gesture	A person lowering her head slowly.	

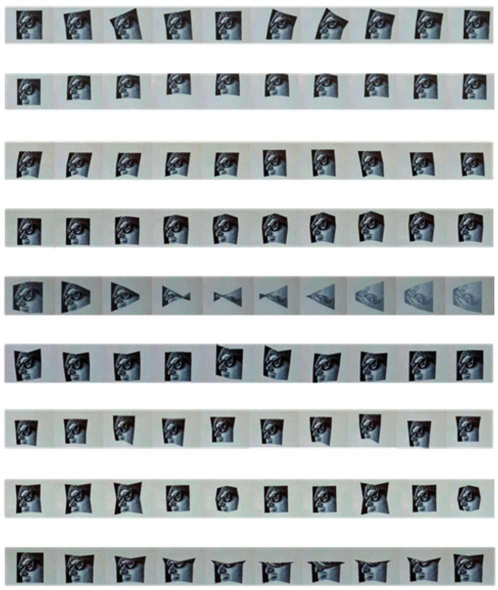


Figure 3.12: The design inspirations of empatheticons and their animation sprites.

overcome limitations of Android Animation API, which only supports affine transformation, such as rotation and scaling. Each picture was divided into a 2x2 grid, which yielded 9 control points. These points served as the boundary of an image and defined the animation of a picture. OpenGL then created a new image by deforming the original one and interpolating colors. The resulting animation is composed of a series of images displaying at 30 frames per second.

Verification of Emotion-Animation Mapping

Before integrating empatheticons into a group recommendation system, we investigated whether people could correctly identify the empatheticons as the emotional states that they intended to represent. We recruited 15 students and researchers, 11 males and 4 females aged between 22 and 33 ($M = 26.8, SD = 1.9$). Participants went over the nine empatheticons one by one, and chose the emotion that they perceived from a given set of options. Results showed that six out of the nine empatheticons were correctly guessed, meaning that the proportion of the option that is supposed to be chosen outweighed the proportion of other options (Table 3.4). Note that Table 3.4 only lists the two most frequently chosen emotions. More specifically, few people mistook *tension* and *sadness*. *Energy* (guessed as *wonder*), *peaceful* (guessed as *nostalgia*), and *tenderness* (guessed as *peaceful*) appeared to be ambiguous. The rest of the empatheticons were correctly recognized only by a small margin.

We revised empatheticon design based on feedback collected from participants. We made

Wonder: wonder = 53.3%; nostalgia = 40%
Transcendence: transcendence = 53.3%; peaceful = 46.7%
Tenderness: peaceful = 46.7%; tenderness = 40%
Nostalgia: nostalgia = 40%; wonder = 33.3%
Peaceful: nostalgia = 60%; tenderness = 26.7%
Energy: wonder = 46.7%; energy = 40%
Tension: tension = 80%; energy = 13.3%
Joyful: joyful = 53.3%; energy = 47.7%
Sadness: sadness = 73.3%; tenderness = 20%

Table 3.4: Percentages of selected labels for empatheticons.

the energy empatheticon more vivid by increasing the intensity of rock gesture and adding short pauses when 'the foot stepped on the ground'. While participants agreed that *peaceful*, *tenderness* and *nostalgia* were difficult to separate and sometimes co-existed, they suggested us to pay more attention to details when illustrating sublime emotions. We thus changed the metaphor of *peaceful* from tranquil water surface to gentle ripples. We also improved *tenderness* by adding feet movement when a person was waving a cradle.

Since we aimed to employ empatheticons in the context of listening to music, we conducted a second round of verification to see to what extent the animations could adequately express the emotions evoked by music, with a group of 42 students (27 males and 15 females) pursuing different levels of educational degrees (bachelor, master or Ph.D.). Five participants were Asian, and the rest were European. None of them have participated in the previous round of verification. This time, a piece of background music ac-companied each of the empatheticons. We asked participants to rate, on a scale of 1 to 5, how well the given empatheticon represented the emotion perceived in the corresponding music. We also encouraged participants to provide comments on the design.

Table 3.5 shows the average acceptance and standard deviations of the empatheticons in representing music-evoked emotions. Most ratings of the empatheticons are above 3.5, indicating that users accepted the empatheticons and believed that they are recognizable. Particularly, *joyful* is considered the most representative ($M = 4.43, SD = .77$). This is mainly due to the vividness of 'jumping' metaphor, as reported by participants. Only the average rating of *transcendence* is below 3.5, with highest standard deviation compared to other empatheticons. A major reason of the lower average rating is reported as users' understanding of the emotion of transcendence per se. As users pointed out in the comments: "*It's difficult to rate this one because I am not very familiar with this word, but I do think the animation matches the feeling expressed in the song very well.*" Since many comments expressed that the empatheticon transcendence can well present the feelings embed in the song ($N = 22$), we decided to accept the current version of *transcendence*.

In summary, participants can well recognize the emotions represented by the empatheti-

Categories	M	SD
Wonder	3.51	1.21
Transcendence	3.32	1.23
Tenderness	3.88	1.04
Nostalgia	3.83	1.10
Peacefulness	4.20	0.95
Energy	4.27	0.87
Joyful	4.43	0.77
Tension	3.98	1.01
Sadness	4.20	0.90
Overall	4.20	0.65

Table 3.5: Average ratings for empatheticons in representing emotions in music context.

cons, especially in the context of music listening. Furthermore, to well present an emotion with an empatheticon, it is essential to choose a representative metaphor and delineating the details of the animation.

Integration of Empatheticons in GroupFun

After verification, we integrated empatheticons in GroupFun. Users can log in to GroupFun with Facebook accounts, create and join a group, invite members, upload music, and listen to a common playlist. They can specify music preference by rating the songs provided by GroupFun. GroupFun will generate music playlists by aggregating group members' ratings. We do not discuss the preference aggregation method since it is out of scope of this paper.

Group music interface contains two sections (see Figure 3.13). Section 1 is a social space, which consists of Facebook profile pictures of all group members (1.a) and emotional states of members that a user chooses to see (1.b). If a member in Area 1.b changed emotion at the timeline of the song, his/her empatheticon updates accordingly. Section 2 is an individual space, including the name and artist of the music, music controller (2.b), current emotion of the user (2.a), music player progress bar (2.c) and emotion selection area (2.d). Users can slide from left to right in Area 2.d to browse all nine emotions. For example, Alicia is listening to the song 'Paradise' by Coldplay. She feels energetic at the beginning, and thus selects *energy* from Area 2.d. She then drags the profile pictures of her friends Lucas and Victoria from Area 1.a to 1.b, from where she noticed that Lucas was feeling *nostalgia* and Victoria was feeling *wonder*. When the song approaches the climax, Alicia feels differently and changes his emotion to *joyful*.

Figure 3.13: Integrating empatheticons to GroupFun.

3.6.2 User Study

To assess the impact of empatheticons in GroupFun, we studied how users may use, perceive, and react to empatheticons when using this group music recommender system. In particular, we investigated *whether* and *how* the group members' emotion could influence a user's behavior.

We decided to conduct a study in our laboratory to make an in-depth observation of the participants and collect their comments when they were using the system. Each participant was invited to the laboratory individually to mimic a non-collocated and asynchronous music-listening scenario. To understand whether and how a user may be influenced by his/her group members' annotations, we preset group members' annotations, which we call **manipulation**. The total number of group members' annotations is set as a controlled variable. The manipulation was not revealed to participants until after the study was completed. This is another reason why we decided to conduct an in-lab study. Both quantitative and qualitative methods were used to analyze the results.

Emotion annotations were logged to analyze the correlation between a user's emotion annotations and his/her group members'. A post-study questionnaire (see Table 3.6) was used to assess user experience such as user-perceived intimacy, immediacy and their satisfaction with the system. The questionnaire was designed based on ResQue model [105], which assesses quality of user experience of a recommender system, and Tu et al.'s measurement of social presence in online learning environment [115]. Each question was responded on a 5-point Likert scale. Pearson correlation analysis was used to investigate the relationship between the survey items and annotations.

Q1: Overall, the emotion interface has successfully visualized the emotions in a musical context.
Q2: The emotion interface is useful in GroupFun.
Q3: The emotion interface is easy to use.
Q4: The emotion interface is easy to learn.
Q5: The emotion interface is novel.
Q6: The emotion interface is entertaining to use.
Q7: I immediately felt my friends' emotions.
Q8: I felt I was listening to music with my friends.
Q9: I paid close attention to my friends' emotions while listening to music.
Q10: My emotion annotations were influenced by those of my friends.
Q11: The songs recommended by GroupFun fit my tastes.
Q12: The songs recommended by GroupFun fit my friends' tastes.
Q13: I am satisfied with group experience with GroupFun.
Q14: I would like to use GroupFun again in the future given the opportunity.

Table 3.6: Survey questions in User Study.

Participant observation method was used to gain an in-depth understanding of how users used empatheticons and their impacts on GroupFun. Participants were encouraged to think aloud during the study. The process during which users interacted with GroupFun was audio-recorded for further analysis. Additionally, each question in the post-study survey was followed by a text area so that the participants could provide comments.

Participants

We recruited 18 participants in 6 groups, each group consisting of 3 members. They were recruited on campus via word-of-mouth. The 6 groups were made of friends, classmates or colleagues. All the participants had Facebook accounts. The participants included 8 females and 10 males from 9 different countries in Europe and Asia (Switzerland, Spain, Korea, China, etc.). Their ages ranged from 18 to 28 ($M = 24.9, SD = 2.3$). All of them were currently using smart phones and were familiar with listening to music on smart phones. At the end of the study, all participants were rewarded with a specially designed music-theme USB stick.

Material

We installed GroupFun on Samsung Galaxy SII 9100 with Android OS 2.3.2. A desktop computer with a 22-inch wide-screen display was used for participants to fill in the post-study questionnaire. We provided them with a list of 9 emotions and explanations on an A4 paper. We also observed user behavior and logged down on paper.

To prepare for the manipulation, we first selected 20 songs from top tracks in Last.fm to cover various genres. The length of the songs ranged from 2.9 to 4.8 minutes ($M = 3.56$). We then preset the emotions for all participants. The preset emotions were collected by inviting two students from the university music association to annotate the songs. We also controlled the number of each song's annotations so that the annotation count of the 20 songs is evenly distributed, i.e., from 5 to 22. We randomly assigned the number of preset annotations of each song by lottery. We then randomly filtered the annotations to meet the controlled number. The categories of preset annotations ($n = 247$) for the 20 songs were distributed as follows: *energy* 11%, *joyful* 18%, *nostalgia* 6%, *peacefulness* 12%, *sadness* 7%, *tenderness* 9%, *tension* 10%, *transcendence* 18% and *wonder* 8%. Finally, we employed another two students to double-check the preset annotations to avoid obvious inappropriate emotions of the songs.

Procedure

Each participant scheduled one hour with us and provided their Facebook homepage via email. We obtained their Facebook IDs from Facebook Graph API and created a group for the participant and his/her two group members in GroupFun. Upon arriving in the laboratory, participants were informed that their group members had participated in the study prior to them, regardless of the actual participation time. For participants who had already known their group members' experiment time, we told them that their friend(s) have rescheduled the time.

Participants were first debriefed about the scenario (same as in Introduction) and procedure of the experiment. They then logged in to GroupFun with their Facebook accounts, entered their group and familiarized themselves with the interface (see Figure 1) using a warm-up song. They practiced to annotate songs with empatheticons and watched their group members' emotions, which were also manipulated. The warming-up session lasted for an average of 2.4 minutes ($Max = 3.2, Min = 1.2$).

Afterwards, they started to listen to 10 songs randomly shuffled from the 20 songs that we had chosen. They could annotate emotions during the timeline using empatheticons. This procedure took up to 40 minutes. Finally, they filled in the post-study questionnaire.

At the end of experiment, we informed the participants of the manipulation in the study, including the fact of presetting emotion annotations and their group members' actual participation time.

3.6.3 Findings

Overall, users found empatheticons easy to use ($M = 4.65, SD = .61$), easy to learn ($M = 4.59, SD = .94$) and entertaining to use ($M = 4.24, SD = .66$). Users also considered empatheticons as a useful tool in GroupFun ($M = 4.18, SD = .72$).

This subsection reports user evaluation of empatheticon interface and how empatheticons affected user-perceived immediacy, intimacy, group dynamics in GroupFun.

Perceived Immediacy

We first assessed the immediacy of empatheticons. By immediacy, we refer to the system's ability to allow users to instantly feel each other's emotions and presence. According to [131], an individual's sense of being with others and immediacy of social interaction are psychological phenomena. User report is conventionally used as a subjective measurement.

User-perceived immediacy is evaluated by Q7 – "I immediately felt my friends' emotions." ($M = 4.18, SD = .53$). User perceived social presence is assessed by Q8 – "I felt I was listening to music with my friends" ($M = 4.24, SD = 1.03$). The results indicate that users could immediately feel group members' emotions and effectively perceive members' social presence, which we refer to as perceived togetherness. Furthermore, eight users have elaborated on perceived togetherness in their comments. For example, User 7 remarked,

> *"I really like to see what my friends' feeling at the moment I was listening, made me feel like we were listening to music together, have more fun."*

These comments further provide evidence that empatheticons could enhance user-perceived togetherness.

By observing users, we also found cues of user enjoyment brought by perceived immediacy and togetherness. All users laughed when they saw their friends' profile pictures jumping and dancing. Some users (User 2 and 18) even stood up and danced together with the empatheticons. The excitement was more obvious when the participant changed the emotion concurrently with her friends. User 16 laughed loudly and showed a surprising face when she changed to joyful simultaneously with her friend. As she said:

> *"It's impressive that we start to change all at once when the song reaches climax."*

The excitement also came when all members in a group were exhibiting the same emotion and movement. As User 12 remarked,

> *"It is exciting to see all of us jumping together. It looks like a group-dance."*

The above observation further explains the statistical correlation between user-perceived togetherness and entertaining of use.

Design Principle 3.6.1. *Consider temporal cues in emotional interfaces to enhance users' perception of immediacy.*

Perceived Intimacy

We also evaluated whether empatheticons could improve intimacy. By intimacy, we refer to the system's ability to bring familiarity and friendship to its members. Similar to immediacy, intimacy is also a subjective metric [131] and we evaluated it through self-report.

Users were asked to rate to what extent they paid close attention to their friends' emotional responses (Q9). Results ($M = 4.53, SD = .62$) indicated participants had attentively watched their group members' feelings during the music timeline. We also discovered some patterns of how users paid attention to group members' emotions. First, users showed interests in comparing their emotions with their friends'. As User 10 commented,

> *"Every time he changed, I started to think why he changed, am I feeling the same, and should I change?"*

At a point when both friends changed to the same emotion, User 8 naturally asked whether he should also change accordingly. Users also expressed doubt when what they felt was obviously different from other members' emotions. As User 2 remarked,

> *"I really want to ask him how he can stay calm in such a joyful song!"*

Additionally, users also showed excitement and pride when they selected an emotion which was immediately followed by other group members (User 18). Furthermore, users showed high enthusiasm in understanding group members' music tastes. For example, User 14 reported proudly,

> *"I think I have found their pattern: less active at the beginning and gradually becoming active. From the frequency of emotion changes, I can guess how much they like the song."*

Users also pointed out empatheticons' potential in enhancing friendship. This mainly refers to the tendency to share emotions and interact more with group members.

> *"I'm more motivated to interact with them. No need to input any text (yeah!!), but using a very cute way to dance together with them. I really feel we were spending time together,"*

commented by User 3. His group mate, User 2, also commented,

> *"Looking at my boyfriend's profile picture soon brings me to think I'm spending time with him. Maybe he is trying to pass on messages in the music."*

Design Principle 3.6.2. *Consider eliciting and sharing users' emotions towards the same items to enhance users' perception of intimacy.*

Group Dynamics

Furthermore, we studied group dynamics by examining whether group influence exists when using empatheticons in GroupFun. By group influence, we refer to members' impact on each other. We investigated group influence by both subjective metric and annotation logs.

Perceived group influence was subjectively measured by Q10 of post-study questionnaire, which asked users to rate to what extent their friends' emotions had influenced theirs. Participants' answers have a high deviation ($M = 3.53, SD = 1.07$). The distribution of the rating is also scattered, with p(2) = 17%, p(3) = 35%, p(4) = 24%, p(5) = 24%, which shows diverse levels of perceived group influence. Users' different attitudes were also reflected in their comments. For example, User 2 showed her willingness to be consistent with other members in her comments,

> *"It's more interesting to see everyone in the group moving and dancing together in the same style. That makes the experience more harmonious and brings more fun!"*

Some participants felt satisfied when their feelings gradually became closer to their friends'. As User 13 commented while listening to music,

> *"I don't want to behave too differently from other members. It might be embarrassing."*

Some participants were unintentionally influenced by members. For example, User 6 reported,

> *"When I saw my friends changing emotions, I naturally wanted to change mine."*

By contrast, some participants were less willing to be influenced by others. User 4 is a typical user who wanted to be unique in a group,

> *" don't want to appear the same as them. Imagine you were in a party. How boring it would be if everyone is the same!"*

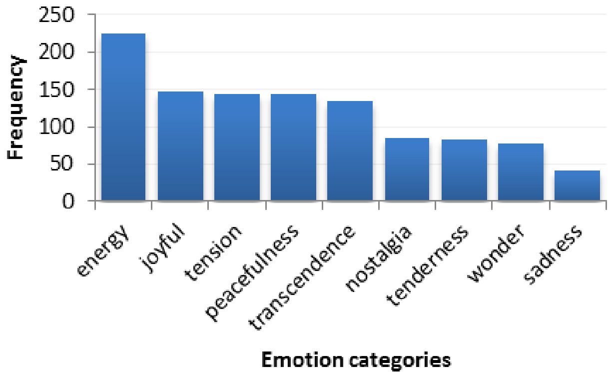


Figure 3.14: Implementation Design.

User 16 also mentioned that he was hardly influenced by others,

> *"For me, listening to music is a conversion with myself. I would rather listen to my heart how I felt about the music."*

Therefore, while empatheticons help us to discover group influence in GroupFun, whether users are actually influenced by group members' emotions may differ from individual to individual.

We also found cues of group influence by analyzing annotation logs. We first present a descriptive statistical analysis of annotations. We have collected a total of 1081 annotations from the 18 users. Users have provided an average of 60.1 annotations using empatheticons ($Max = 112, Min = 38, SD = 5.04$). The distribution of annotations by emotion is shown in Figure 3.14. The most frequent emotions are energy (n = 225, p = 21%) and the least frequent one is *sadness* (n = 43, p = 4%).

We then calculated each user's emotion agreement, defined as the case when the selected emotion is the same with at least one group member at the time of annotating. Dividing by the user's total annotation count, we calculated the ***agreement rate***, with the highest at 50.0% and lowest at 8.3%. Admittedly, the agreement could also result from emotions evoked by music per se.

We further examined the impact of group members' activeness in annotation. We measured *activeness* using the number of annotations. Since group members' annotations were actually manipulated, we evaluated the relationship between the number of user annotations and preset annotations. Figure 3.15 shows the distribution of preset annotation count and the average annotation count (with minimum and maximum) for each song. By preset annotation count, we mean the sum of two group members' annotations. A two-tailed Pearson correlation test showed an approximately significant correlation ($r = .425, p = .079$) between the preset annotation count and users' average annotation count of each song. It indicates that the more

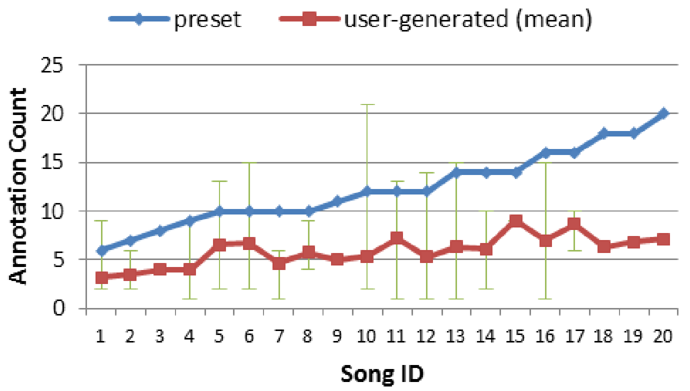


Figure 3.15: Implementation Design.

active the group members are, the more annotations a user is likely to provide.

Design Principle 3.6.3. *Consider creating positive group influence and emotional contagion to enhance group consensus.*

Group Satisfaction

We also evaluated users' group satisfaction with GroupFun ($M = 4.47, SD = .623$). A paired-samples t-test showed that the rating is significantly ($p < .01$) higher than that of whether the songs match their tastes ($M = 3.59, SD = .618$). This implies that empatheticons can effectively enhance group satisfaction compared to their individual preference of the recommended songs.

Furthermore, users' overall satisfaction is highly correlated with their number of annotations (r = .684, p = .002). This indicates that the more active a user is in annotating emotions, the more likely s/he is satisfied with GroupFun. In other words, the total number of annotations could be an indicator of a user's satisfaction with the system. As User 17 commented,

> *"I not only focused on enjoying the songs, but also how others feel. They were communicating their feelings with me which motivated me to express myself. This helps me to understand and respect their preferences. Not only my friends' opinions are considered, but I learn to appreciate their favorite songs."*

This indicates that sharing emotions and learning each other's preferences help users to accept each other's music taste and enjoy the music. Group satisfaction is also significantly correlated with perceived entertaining to use ($r = .565, p = .018$) and perceived togetherness ($r = .496, p = .043$). Thus creating co-presence feature and making it entertaining to use can enhance users' satisfaction of a group recommender. Users also mentioned what made them consider using GroupFun again. For example, User 2 commented,

> *"I really think this kind of software can have an effect on the relationship on the people in the same group. What can make the relationship stronger? I think it is sharing: happiness, sadness and those crazy moments. And it also helps us understand other people more."*

This hints that designing for features that help users enhance social relationship (i.e., intimacy) could promote user retention of a group recommender.

Design Principle 3.6.4. *Consider designing social interfaces that enhance group members' intimacy and immediacy during the interaction to enhance their group satisfaction and intention to use the system.*

3.6.4 Summary

In a group recommender system, mutual awareness is crucial for users to gain a better understanding of each other's preferences. It is also an essential element in maximizing group satisfaction. To achieve mutual awareness, prior studies have presented solutions in membership awareness, preference awareness and decision awareness. In this work, we investigated emotion awareness and its impact in promoting group satisfaction. We first designed a set of interface elements called empatheticons. These empatheticons were verified and improved during two initial verification studies, showing that users could easily map the empatheticons to their respective emotions in a musical context. The empatheticons were then integrated in an empirical environment – GroupFun – to validate our hypotheses. In an in-depth user study, we showed that empatheticons could effectively enhance users' perceptions of togetherness, i.e., the feeling of listening to music together with friends. Additionally, empatheticons also served as a useful tool with which users can provide emotional annotations and familiarize themselves with group members' preferences. Furthermore, the number of a user's annotations was found to be influenced by other members' and enable positive group dynamics.

A few limitations are worth mentioning. Firstly, our study was conducted in the laboratory with manipulation. An in-situ study can bring us with further insights into the use of empatheticons. Additionally, we investigated GroupFun with empatheticons to obtain an in-depth understanding of users' attitudes and behavior patterns. In the future, we will compare our findings to GroupFun without empatheticons.

3.7 Chapter Summary

In this chapter, we study social interface design for group recommender systems. We first developed an experimental platform called *GroupFun* – a music group recommender system with emphasis on mutual awareness features. Realizing the lack of study on emotion awareness in group recommenders, we design three types of emotion representations – emotion wheels (for *ACTI* and *CoFeel*) and dynamic animations (for *empatheticons*). We integrate these interfaces into GroupFun and iteratively improve our design by evaluating with a total of 30 users. We then report an in-depth laboratory study to assess the effectiveness of empatheticons and the roles of emotion awareness tools in GroupFun. The design and evaluation process of these emotion awareness tools progressively lead us to a number of important findings and design inspirations. First of all, emotions, which *seem* to be irrelevant with th recommendation process, turns out to mediate interpersonal relationship within an online group and implicitly help group members influence each other and reach better consensus. Additionally, visualizing emotions with temporal cues increases users' feeling of connectedness and perceive the interaction as real-time even in an asynchronous communication. Studying through the lens of group recommender systems, we aim to evoke researchers to contemplate emotions – the easily-neglected yet potentially-meaningful information in the broader areas of groupware.

4 Social Interaction Design for Group Recommenders

In this chapter, we investigates how to design social interactions to motivate users in persuasive technologies. More concretely, we focus on social interaction mechanisms in pervasive fitness applications. Part of this chapter is published in [37, 38].

4.1 Introduction

Citizens of the industrialized nations are susceptible to inactivity and stress due to sedentary and highly competitive lifestyles, which can lead to serious health problems such as cardiovascular diseases, diabetic type II, depression, and cancer. The World Health Organization defines *health* as a “state of complex physical, mental and social well-being and not merely the absence of disease or infirmity." Both research communities and commercial sectors are putting an increasing effort to develop wearable sensors and mobile applications that help and “nudge" individuals to increase their physical activities, eat a healthier diet, better manage their sleep and stress, and engage in social lives with family and friends.

Wellness and lifestyle management systems using wearable sensors and mobile apps are on the rise [128]. Many of these applications use gamification to motivate users to exercise more. Gamification is the use of game elements in non-game context in order to engage users [51]. Traditional methods have focused on competition such as leaderboard for community users [21], self-reflection such as metaphorical visualization of garden flowers for individual users [55, 83, 84], or a combination of the two [85]. Recent work has identified crucial evidence supporting social interaction as an essential element to motivate users to perform physical activities in designing pervasive fitness apps [29, 41]. Such social interaction schemes are called ***social incentives***. In particular, such social interaction among group members includes sharing physical activities, cooperating and competing. However, a detailed examination of users’ behavior in a group environment remains an open subject of study.

In this chapter, we present the design and implementation of a mobile social game called **HealthyTogether**. We further employ HealthyTogether as an experimental platform to investigate various social incentives. We then report the results of three comparative user studies with HealthyTogether that explore the effectiveness of *competition*, *cooperation*, *social accountability* and *social capital* as incentives in persuasive technologies for physical activities.

4.2 Related Work

Gamification component is frequently used in designing persuasive technologies to make sports activities fun. Methods such as leaderboards and metaphorical visualization have been used to encourage individuals to work out more. Most methods rely on either community or self-perception for the implementation of the games. Many commercial fitness products, such as Nike+ [8], Fitbit [5], miCoach [7], adopt community-based competition. A number of systems also use metaphors to present physical activities. UbiFit Garden [42] is a mobile application that visualizes users' daily steps by the growing status of plants. The more activities a user takes, the healthier his/her plant looks. UbiFit Garden also encourages users to take various types of activities by displaying butterflies. Fish'n'Steps [85] uses the metaphor of fish tanks to visualize the step count of users. Using a metaphor connects the user's physical activities with a living creature. The empathy to take care of a plant or animal is shown to motivate a user to take exercise. Fitster [21] accumulates users' exercise distances and translates them to the routes on geographical maps, so that users can treat their exercises as traveling around cities. The system iFitQuest [87] uses several metaphors, such as escaping the wolf and collecting the coins to enhance different aspects of physical performance: speed, endurance, etc. The above work mainly motivates users in an individual setting.

Social accountability has also been shown to be effective in helping users to achieve goals. Ahtinen et al. [20] have found out that connecting with family members and loved ones can help motivate users; connecting with people with similar wellness targets from communities within short distances can also increase motivation towards wellness activities. Stickk.com [13] is a website that helps users to achieve their goals by allowing them to appoint another person to monitor the progress and verifying the accuracy of progress report. They can add supporters who can encourage or cheer them by putting comments on their progress journal. Users can also put stake on the goal and specify where the stake would go if they fail in the goal. GoalSponsor [6] is a mobile application that allows users to set up goals and sponsors whom they should be accounted for. A sponsor can be a friend, a professional in healthcare, or someone who has accomplished the goal successfully. Every day, users check in to report whether they have made the effort for the goals and upload pictures as evidence. Users are more committed in fulfilling the goals either because they do not want to let others down or because they do not want to lose reputation in front of others. In the above work, the structure includes one person who has a goal to fulfill and another person who is monitoring and supervising on the progress. However, little work has investigated the mutual accountability between exercising buddies.

Many applications allow users to form a team and compete with each other. Fitster [21] visualizes users' steps in a social group and motivates users through a virtual competition environment. Kukini [29], Fish'n'Steps [85] and Life Coaching Application [59] support competition by helping users to form a team and explicitly introducing social interaction and social pressure. Cooperation, which binds users' performance with that of their team members, not only promotes users' physical activities, but also brings social benefits such as group enjoyment and socialization with friends [95]. In Fish'n'Steps, any insufficient performance of group members deteriorates the environment of the fish tank, such as darker water, removal of tank decoration, etc. Users are thus motivated to enhance their performance for the sake of group responsibility. Chick Clique [112] calculates and displays the average steps of a team. This makes users aware of which members are below average and stimulates them to encourage teammates who are left behind.

Several studies have documented users' qualitative feedback about competition and cooperation in mobile fitness applications. For example, Lin et al. [85] found that some users considered competition engaging while others believed it was unnecessary and incompatible with the theme of Fish'n'Steps. They also discovered that cooperation did not produce any significant improvements if team members were anonymous. Qualitative results of Ahtinen et al. [20] indicated that users were motivated by workout as a group, which they referred to as cooperation. As for competition, users would rather compare with themselves than competing with others. Macvean [87] demonstrated that users' background should be considered carefully in competition setting. Findings of the above studies were based on evidence collected by users' qualitative feedback about social incentives. A controlled study comparing competition and cooperation is lacking. Halko and Kientz [64] found, through large-scale online survey, that users of conscientious personality traits are more likely to use both competition and cooperation strategies and that people of agreeableness and openness traits tend to favor competition. However, their study focuses on users' proposition to social incentives and the results are not validated empirically. The closest one to our study is conducted by Peng and Hsieh [101], who found that cooperative goal structure lead to greater effort put into motor game than the competitive goal structure through a controlled user study. However, they did not generalize their findings in mobile fitness apps.

In order to compare the various social interaction incentives, we designed and developed an experimental platform based on literature that guides fitness application design. Consolvo et al. [41] have proposed the following design requirements for technologies that encourage physical activity: 1) give user proper credit for activities, 2) provide personal awareness of activity level, 3) support social influence, and 4) consider the practical constraints of users' lifestyle. We emphasize on the following aspects in designing HealthyTogether. First, HealthyTogether embodies different social settings in the rules of the game and reward users for their performance in the group. Second, HealthyTogether visualizes information that provides mutual awareness for exercise buddies. Third, it displays activity information of both users and their buddies. Fourth, it retrieves users' activity data from an off-the-shelf sensor that could conveniently fit into users' daily life. We describe the design and implementation of HealthyTogether in the next sections.

In our work, we aim to integrate various social interaction mechanisms in HealthyTogether and compare their impact. At the meantime, we have improved the interface of HealthyTogether after each user study based on user feedback. Thus, we refer to the different versions of HealthyTogether as v1.0, v2.0 and v3.0.

4.3 Social Accountability as Social Incentives

In this section, we investigate social accountability as social incentives in persuasive technologies. We report the design and results of a user study involving 24 participants for a period of two weeks, using HealthyTogether v1.0 as an experimental platform. We then derive design guidelines based on our findings. In the rest of the section, HealthyTogether refers to the version 1.0 by default.

4.3.1 Experimental Platform: HealthyTogether v1.0

We started investigating various social gamification mechanisms by examining their effects on two-person groups, namely dyads. To this end, we developed a mobile application on Android platform called HealthyTogether that involves a pair of users to exercise together and earn badges based on their walking and climbing performance.

To measure users' activities, we chose Fitbit among a number of off-the-shelf activity trackers. First, it is unobtrusive and convenient to wear. Second, compared to other competitor products, Fitbit (as shown in Figure 4.1a) and 4.1b)) is a comprehensive tracker that records not only steps and distance, but also floors and sleep. Last but not least, its API allows us to interact with Fitbit data in HealthyTogether. HealthyTogether automatically retrieves users' step and floor information from Fitbit server every two minutes. It also supports manual synchronization.

Game Rules

We designed a series of rewarding mechanisms for HealthyTogether in order to investigate the impacts of different social settings in pervasive fitness application. A user can win badges based on Karma Points, which are calculated as below.

$$kp_{step}(u) = a \times step(u) + b \times step(u'),$$

where u stands for a user and u' stands for his/her buddy. Based on different values of a and b, HealthyTogether provides the following three social settings.

- **Competition setting**, where a = 100%, b = 0;
- **Hybrid setting**, where a = 80%, b = 20%;

Figure 4.1: HealthyTogether main interfaces and Fitbit tracker interfaces: a) The Fitbit tracker, b) Fitbit in use, and c) the Android phone.

- **Accountability setting**, where a = 0, b = 100%.

In **competition** setting, a user's Karma points are calculated purely by his or her steps. To gain more badges, a user only needs to focus on his or her own activities even if he is exercising with a buddy. Thus, we name this rule competition setting.

In **accountability** setting, a user's Karma points are calculated by the steps of the buddy. Therefore, the more he encourages his buddy to exercise, the more points he earns. Thus, we name it accountability setting. On the other hand, even if a user does not move at all, he can still gain badges from the buddy's activities.

In the **hybrid** setting, a user's Karma points are calculated based on both his (her) own and that of the buddy, proportionally. The idea behind this reward scheme is to encourage competition while also motivating users to cheer each other. Initially, we set *a* and *b* based on the well-known Pareto Principle [10]. In the future, we will also experiment different ratios of competition and social accountability, such as 50%-50% and 20%-80%.

Badges

HealthyTogether issues badges based on $kp_{step}(u)$. The first badge is issued if $kp_{step}(u) > 500$. We choose 500 to help users get started in a short time. This number is followed by 1000 and 2000 and then increases by every 2000 points.

HealthyTogether calculates Karma points on a daily basis but accumulates badges over time. For example, if a user earns 5,353 Karma points in a day, he can gain 4 badges, i.e., 500, 1000, 2000 and 4000. If a user earns 5353 and 6086 points in the first two days, he can gain 4 and 5 badges respectively and a total of 9 badges.

Social Interface

The main interface of the HealthyTogether system is shown in Figure 4.1 c). It contains a 'self' tab and a 'buddy' tab. Each tab displays information about step count, active time and badges of the current day. We use a pie chart to visualize the proportion of time that a user is in various activity modes, i.e., sitting, lightly active, fairly active and very active.

The badge area displays the total number and the various types of badges that a user has earned. The badges are accumulated over time. In Figure 4.1c), the user has earned 6 types of badges with a total number of 16. When he/she clicks on a badge icon, a dialog box will pop out explaining the details of this badge type, including how many badges the user has earned in this type and how he/she earned the badges (see Figure 4.2).

There is a messaging button on the top-right corner of each page. When it is clicked, users can either view message history (Figure 4.3 a)) or send messages to their buddies (Figure 4.3 b)). Users will receive a vibrated notification when buddies send them new messages.

4.3.2 User Study

To study different game settings in real situations, we designed an exploratory deployment study. We first conducted a user study (Study 1) that spans for six continuous working days, which was divided into a three-day control session and a three-day experimental session. After conducting the study, we were able to discover some interesting results. For example, participants suggested that we extend the study to two weeks, excluding the weekends, so that the control and the experimental sessions span over identical days of the week, thus minimizing the influence of a given day's schedule to the physical activities being monitored. For example, a user may work in the office on Mondays but conduct experiments in the laboratory on Wednesdays. We therefore conducted the second study (we name it Study 2) with duration of two weeks. We refer to the control session as Phase I and experiment session as Phase II in both studies.

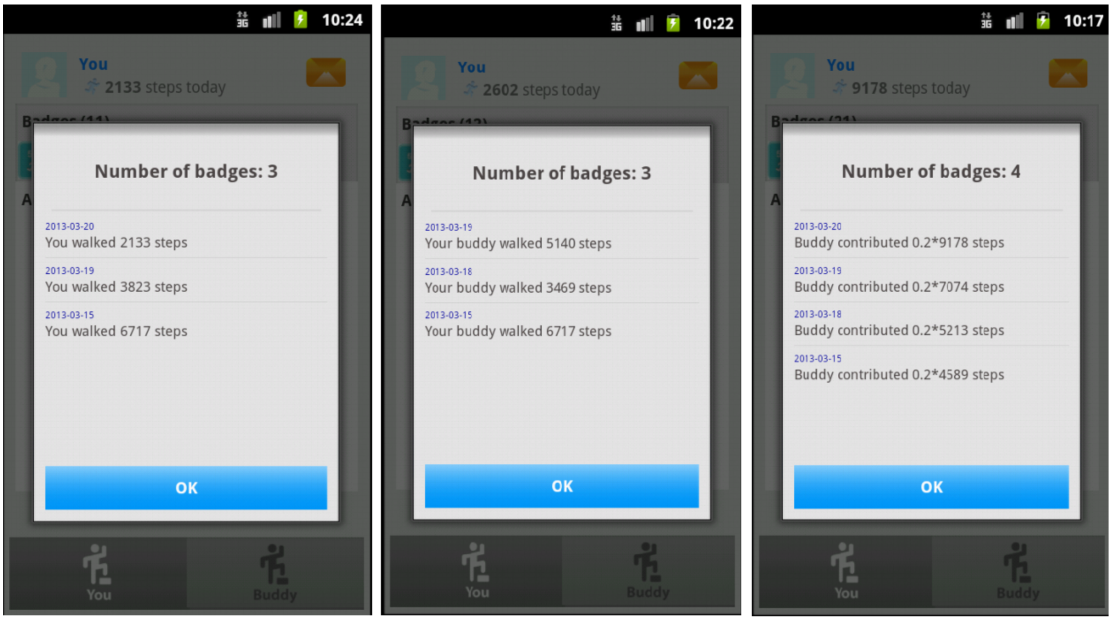


Figure 4.2: Explanation on reasons why users earn a badge. Left: competition setting; middle: social accountability setting; right: hybrid setting.

Participants

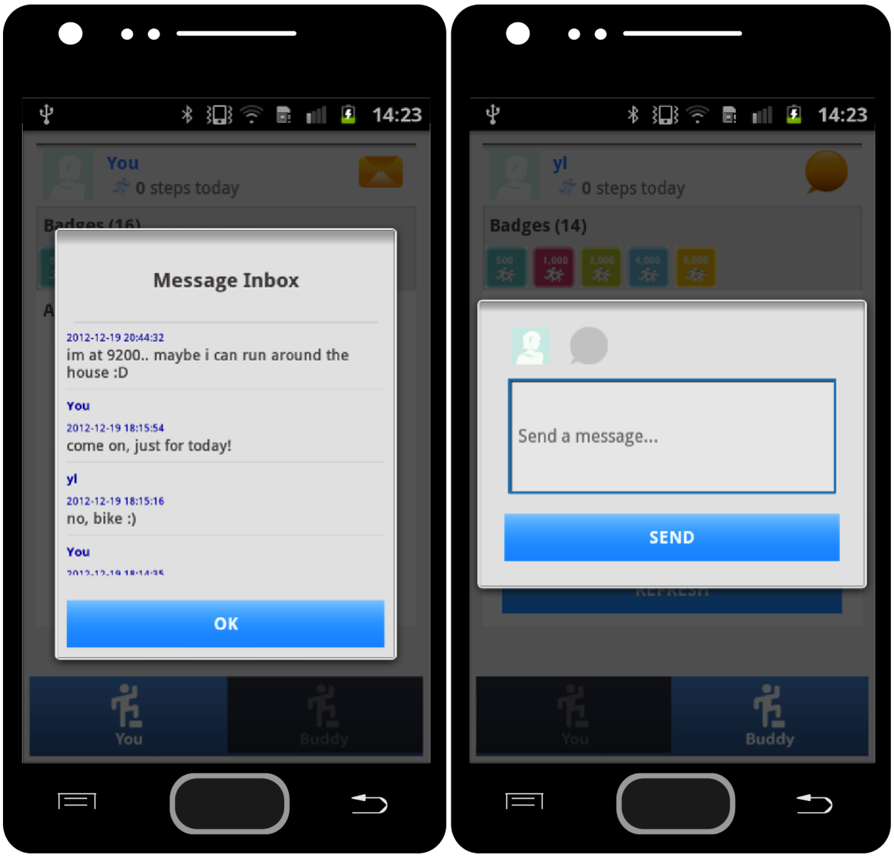
We recruited the participants on campus via word-of-mouth. After one person signs up, we asked her to invite a buddy of her choice to join. We required that each dyad should not work in the same office or too close to each other. We offered all participants a 50-CHF gift card as compensation for their time. We mainly recruited two types of buddies (i.e., friends and colleagues), as shown in Table 4.1. None of them have used Fitbit before.

Materials

We provided each user with an Android phone with 3G SIM card and a Fitbit. Three users requested to use their own Android phones because it would be more convenient for them. We checked that their phones were compatible for installing HealthyTogether.

Procedure

Both Study 1 and Study 2 were structured as a two-phase, within-subjects design. *Phase I* allowed participants to become accustomed to using Fitbit and allowed us to collect baseline fitness data. In this phase, all participants used Fitbit alone without connecting with buddies. In *Phase II*, participants in baseline groups (Group A1 – A3) continued to use only the Fitbit while groups in social settings (Group B1-B3 in competition setting, C1 – C3 in accountability setting, D1 – D3 in hybrid setting) started to use Fitbit and HealthyTogether with buddies. The structure of Study 1 was the same as the Study 2, except the duration was extended to two



a) b)

Figure 4.3: Messaging functions in HealthyTogether.

weeks, with both phases extended from three days to five days.

At the beginning of the study, we invited each pair of participants to our laboratory and helped them to set up their Fitbit accounts. We also interviewed them on their experience in using fitness sensors or devices. At the end of Phase I, we invited participants in social settings to our laboratory again to install HealthyTogether with different game rules.

Since our user study last for up to two weeks, we requested participants to fill in a daily experience survey related with the study. At the end of each day, we sent a reminder email with the survey link to participants asking them whether they have anything to share with us about their experience using Fitbit or HealthyTogether. The survey only contains one question: "Do you have anything to share with us on your experience using Fitbit/HealthyTogether today?" Daily survey not only helps us to gain an in-depth understanding of users' experience, but also facilitates us to explain their step data with activities during that day.

At the end of the study, we organized a semi-structured interview. We invited both participants of each dyad to attend the session together, so that they could share their stories. We did

Dyad ID	Type	Location	Phase II (Team)	Ages
Study 1	A1	Friends	Different buildings	28 to 30
	B1	Friends	Different buildings	25 to 28
	C1	Friends	Different buildings	22 to 24
	D1	Colleagues	Different offices	27 to 30
Study 2	A2	Friends	Different offices	24 to 33
	A3	Friends	Different buildings	25 to 26
	B2	Friends	Different buildings	22 to 24
	B3	Colleagues	Different offices	27 to 33
	C2	Friends	Different buildings	20 to 25
	C3	Colleagues	Different offices	26 to 28
	D2	Colleagues	Different offices	26 to 27
	D3	Friends	Different buildings	25 to 26

Table 4.1: Summary of participating dyads in Study 1&2.

not ask a fixed set of questions, but mediated the session with the aspects their experience and attitudes using HealthyTogether.

4.3.3 Findings

In this section, we report both quantitative and qualitative results collected in Study 1 and Study 2. To facilitate describing results, we encoded the two participants in each dyad with ‘a’ and ‘b’ together with their group ID. For example, we encoded the two participants in Group C1 as ‘C1a’ and ‘C1b’ respectively.

Quantitative Results

Study 1 We first investigate users’ step count across the 6 days. The overall average daily step count is 7439 (min = 3185, max = 11490). We then compare the average daily step count between baseline group (A1) and groups who used HealthyTogether (B1 – D1) to evaluate the effectiveness of social interaction incentives (see Figure 4.4). Results show a slight decrease of steps from Phase I to Phase II across all groups. One explanation is the novelty effect of using Fitbit in the first 1 – 2 days, as reflected in daily survey. One interesting finding is that in Group A1 the average step count decreased by 20.4% but in Group B1—-D1 it decreased by only 10.6%. This implies that **HealthyTogether with social settings could help users to persist in physical activities**.

We further compare Group B1 – D1 with different rules of calculating Karma points. Results show that users in Group B1 (competition setting) have largest difference of Karma points (Δkp = 14556) compared with Group C1 (Δkp = 839) and Group D1 (Δkp = 2982). One explanation is that participants in Group B1 focus more on their own performance compared with other

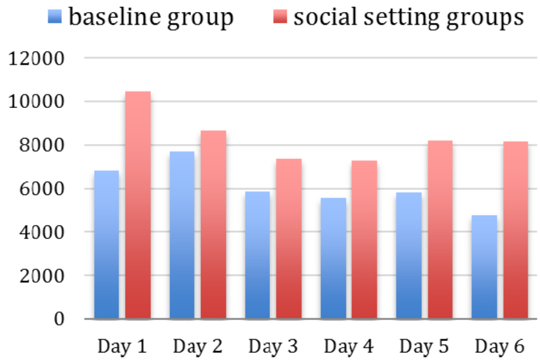


Figure 4.4: Distribution of average daily steps for groups with non-social setting vs. social setting in Study 1.

groups. In other words, it implies that **the social accountability factor, which is applied in Group C1 and D1, could lead to more balanced performance of physical activities between buddies**.

Meanwhile, we have collected a total of 72 messages, shared by Group B1 ($N = 43$), Group C1 ($N = 27$) and Group D1 ($N = 2$). The distribution shows that Group C1 (accountability setting) and D1 (hybrid setting) interacted more compared with Group B1 (competition setting). Particularly, participants in Group D1 exchanged 59.2% more messages than Group C1. This implies that **hybrid setting is most useful to encourage participants to interact with each other**.

Several topics emerged from the analysis of message content, and we present them with sample text in Table 4.2. The distribution of each topic shared in Group B1 – D1 is shown in Figure 4.5. It reveals the following phenomenon: 1) in total, there are 27 chat messages, which have the largest share; 2) encouragement is the main topic that is relevant with physical activity (20 messages); 3) Group C1 has the largest share (13 messages) in encouraging messages; 4) the major topic of Group D1 is workout together (16 messages), and it only appears in Group D1. The results imply that **hybrid setting introduces most conversation in the topic of workout together**.

Topics	Examples
Self-reflection	"*im at 9200.. maybe i can run more*"
Encouragement	"*you should make it 8k for a new badge*"
Comparison	"*the first time i am higher than you!!!*"
Workout together	"*we should walk around the floor together to take a break:)*"
Chat	"*feeling so tired now, go to bed soon*"

Table 4.2: Message topics and examples in Study 1.

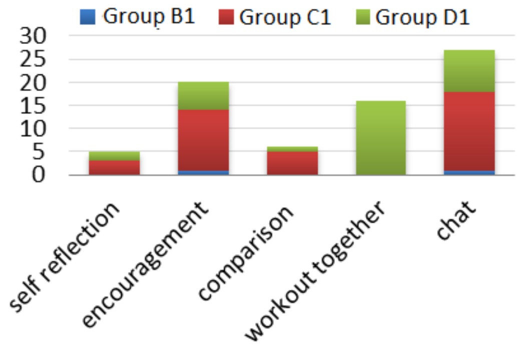


Figure 4.5: Topic distribution of messages exchanged using HealthyTogether in Study 1.

Study 2 The deployment of Study 2 is the same as Study 1 except for the duration of the study. We first verify whether discoveries we found in Study 1 still exist in Study 2. The overall average daily step count is 9501 (Min = 3200, Max = 24334). Figure 4.6 is a distribution of average daily steps between baseline groups (Group A2, A3) and social setting groups (Group B2, B3, C2, C3, D2, D3) in two weeks. The distribution shows a steady increase of average daily steps in social groups from Phase I to Phase II. Comparing social setting groups and baseline groups, we found average steps increased by 9.8% from Phase I to Phase II in social setting groups but decreased by 10.1% in baseline groups. This is consistent with implication in Study 1 that **social settings could motivate users to exercise compared to when they walk alone**.

We then compare groups using HealthyTogether in different social settings. Figure 4.7 shows each participant's average daily steps in Phase I vs. Phase II. The average daily step in competition groups (B2 and B3) have increased from 9747 to 10128 ($\Delta step$=381). In accountability groups (C2 and C3), this number increased from 8888 to 9717 ($\Delta step$=829). Participants in hybrid setting (D2 and D3) have walked an average of 12437 steps per day, with an increase of 1675 steps compared to when they walked alone. The average daily step increase of hybrid group is 51% more than that of accountability group and three times more than that of competition group. If we assume that participants have the same schedule of the same workday in different weeks, and that Phase I is a baseline for participants, then the above results suggest that **hybrid setting encourages users to walk more**.

We further examined the steps between two participants in a dyad. From Figure 4.7, we can see that dyads from both accountability groups (C2 and C3) and hybrid groups (D2 and D3) have increased steps together in Phase II compared to Phase I. On the other hand, the distribution shows that in competition groups, average step count of B2a and B3a have increased 22.5% and 11.0% respectively from Phase I to Phase II while this number decreased for their buddies (2.1% for B2b and 17.8% for B3b). This implies that **accountability factor helps dyads to walk more steps together**. In Study 1, we found the tendency of participants in accountability group and hybrid group to achieve a balanced number of badges. Even though we do not have the same finding in Study 2, the results concur with the implication in

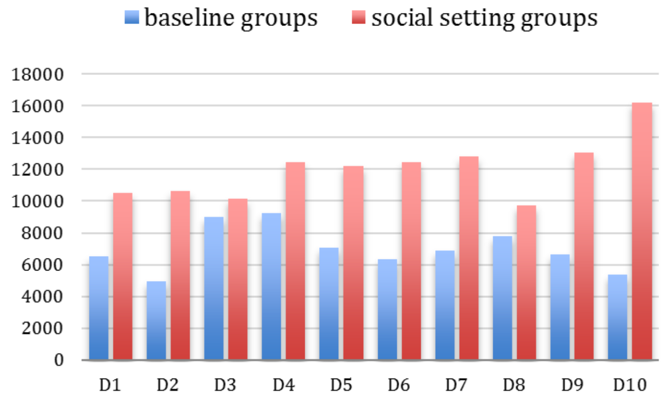


Figure 4.6: Distribution of average steps non-social setting vs. social setting in Study 2 within 10 days.

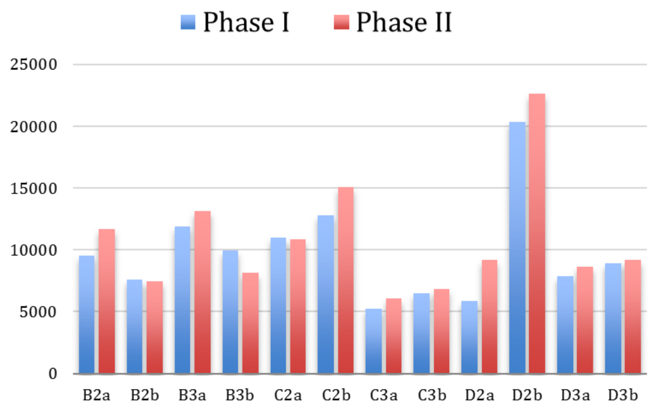


Figure 4.7: Average step count Phase I vs. Phase II in Study 2.

Study 1 that **users have more balanced working performance in the sense that both users in a dyad improve together**.

We then analyze the 86 messages sent between the dyads in Study 2. Participants in hybrid groups sent 58 messages, which is more than twice the number of accountability group ($N = 21$) and seven times more than that of competition group ($N = 7$). Figure 4.8 shows the distribution of message topics within groups of the three social settings. Different from Study 1, messages with topics about self-reflection and encouragement have the largest share in the total of messages (27.8% for both topics). We also discover that hybrid groups have the largest share of messages (81%) in the topic of self-reflection. The distribution accords with what we have found from Study 1 that 1) hybrid groups have most share of messages (75%) in the topic of workout together, and 2) encouragement is the major topic (54%) in accountability groups. If we consider the number of messages as one metric to evaluate social interaction, results in Study 2 provide further evidence that **hybrid setting is more likely to stimulate social interactions**.

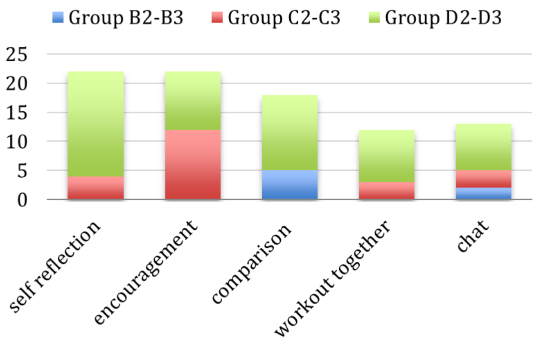


Figure 4.8: Topic distribution of messages exchanged using HealthyTogether in Study 2.

Qualitative Results

In this section, we report the results of daily survey and post-study interviews from both Study 1 and Study 2. During the post-study interview, we found some qualitative results related to our findings from quantitative analysis.

Overall, the feedback about HealthyTogether is very positive. First, HealthyTogether has helped them to compare with each other. As B1b said in the interview,

> *"... to check her (buddy's) steps and compare with mine is most important for me...."*

Second, they could interact with each other via HealthyTogether. B1a reported in her survey:

> *"I received message on the first day using HealthyTogether, so the next day I intentionally walked between the buildings (to gain more steps)."*

Design Principle 4.3.1. *Design social interfaces that allow users to share physical activity data within groups to help them compare with each other*

We also found some evidence that the accountability factor (applied in accountability and hybrid setting) could help users to care about each other. As C1b reported in her daily survey,

> *"I discovered his (my buddy's) step is twice more than mine. As his badges depend on my steps, I feel I should walk more in order not to discourage him."*

This could explain what we found in Study 1 that the participants in C1 have more balanced performance when using HealthyTogether. Additionally, when we asked whether users about their social relationship before and after using HealthyTogether, participants in D3 revealed that they were already very close friends. Participants in C1, C2, D1, and D2 reported that they had developed further relationship with buddies after using HealthyTogether. For example, D1b said,

> *"Even though we are colleagues, we did not talk much. Finding something to do together rapidly brought us closer."*

For another example, C2a reported that she knows more about her buddy:

> *"When I woke up, my buddy already had 3,000+ steps... She is already on campus..."*

However, we did not find the same report from competition groups. This suggests that **social accountability factor could be helpful for a user to enhance interpersonal relationship with the buddy**.

Design Principle 4.3.2. *Consider applying social accountability to help users to enhance their relationship.*

Participants have reported their concerns regarding competition setting. Both B1a and B1b reported competing with each other caused de-motivation.

> "*I knew I could never win over him because he needs to walk a lot from home to work.*" –B1b
> "*I clearly had advantage in winning the game and sometimes I'm afraid walking too much brings her pressure..."* –B1a.

B3a also reported that when he noticed his buddy walking less, he also became lazier. Admittedly, the de-motivation effect could be caused by the unbalanced abilities in physical exercise. This suggest that **choosing a suitable exercise buddy can be important in order to reduce the de-motivation effect of competition**.

Evidence shows that the hybrid setting is more preferred than the accountability setting.

> "*I feel it is a little bit wired if my badges only depend on my buddy's steps. Is it a little bit de-motivating me?*" – C3a

By comparison, instead of 'depending on' others, both D1 and D3 have reported they arranged activities together, such as going for a coffee together by taking the stairs (D1), walking to a farther away cafeteria to have lunch (D1), and going to Zumba course together (D3). D2a reported in his survey that his buddy gave him suggestions how to increase the steps:

> *"Without trying too hard, I almost reached the daily goal of 10k steps. Seeing (name of buddy)'s progress during the day helped to motivate me to move more. It was also useful to talk to him (via message) - he gets a high step count by walking around the room while reading articles, which is a good idea. I did this walking in circles thing too for about 1h, which helped my brainstorming."*

Even though D2b (M = 20372) have clear advantages over D2a (M = 5852) in Phase I, this discrepancy did not demotivate either of them. Instead, D2a's average daily step has increased by 57.3% (M = 9207) from Phase I to Phase II and D2b has increased by 11.2% (M = 22661). As commented by D1a:

"Helping others to become better is a 'plus' rather than 'minus' to your life."

Design Principle 4.3.3. *When designing social interaction schemes, consider balancing between the benefit of the accountability mechanism that allows users to take care of each other and that of the competition that provides independence.*

4.3.4 Summary

In this work, we have developed a mobile application called HealthyTogether that allows dyads to participate in daily physical exercises as a game. We conducted a user study with 12 dyads with a period of up to two weeks and compared participants using HealthyTogether in three social settings and a baseline non-social setting. Results show that social settings, even in the competition mode, can help users to persist more in physical activities compared with the baseline groups. Furthermore, the hybrid setting is more likely to motivate users to walk more and more actively help others. More concretely, the average daily step increase is 1,675 in the hybrid setting, which is twice the number of step increase in the accountability setting ($\Delta steps$ = 829) and three times of that in the competition setting ($\Delta steps$ = 381). Furthermore, the number of messages sent between participants in the hybrid settings is eight times more than those in the competition setting and twice of those in the accountability setting. Integrating social accountability factor is also promising to enhance social relationship between buddies. Finally, we present three design suggestions for designing social interface that use gamification in pervasive fitness applications:

1. Design social interfaces that allow users to share physical activity data within groups to help them compare with each other;

2. Consider applying social accountability to help users to enhance their relationship;

3. When designing social interaction schemes, consider balancing between the benefit of the accountability mechanism that allows users to take care of each other and that of the competition that provides independence.

In the future, we plan to conduct a larger scale of longitudinal studies using our new design of HealthyTogether with more users in various conditions to validate our findings.

4.4 Competition and Cooperation as Social Incentives

We continue to investigate the effectiveness of competition and cooperation as social incentives for persuasive technologies. Based on the previous user study, we improved the design of HealthyTogether, which we refer to as v2.0. We then used it as an experimental platform to compare competition and cooperation through a user study involving 36 users for two weeks. In this section, we report the findings of the user study and derive design guidelines. In the rest of the section, HealthyTogether refers to the version 2.0.

4.4.1 Experimental Platform: HealthyTogether v2.0

The main interface of HealthyTogether v2.0 consists of two parts: game and log. The game area contains a ‘Steps’ tab (Figure 4.9b)) and a ‘Floors’ tab (Figure 4.9c)), which display the dyad’s performance in taking steps and climbing the stairs, respectively. Users’ effort in taking stairs is counted as the number of floors. In each tab, a progress bar visualizes a user’s current number of steps or floors in orange color, as opposed to that of his/her buddy in blue – the opponent color of orange. This is followed by another progress bar which presents how the dyad’s step or floor count contributes to the badges and how many steps or floors the user needs to gain the next badge. The area between the two progress bars illustrates the game rule, which is explained later. The dyad can send messages to each other by pressing either of the two blue icons: cheering face and taunting face. The subsequent badge area shows badges the user has earned during the day. Moreover, HealthyTogether v2.0 allows users to view their performance history of the recent seven days. In the log area, users can log their mood, social interactions, food intake and activities as well as view log history. We do not present this area in detail, since it is out of the scope of this thesis.

Game Rules

We designed a series of rewarding mechanisms for HealthyTogether v2.0 in order to investigate the impacts of different social settings in persuasive technologies for physical activities. A user can win step and floor badges based on his/her Karma points, which are calculated as following.

$$kp_{step}(u) = a \times step(u) + b \times step(u')$$

$$kp_{floor}(u) = a \times floor(u) + b \times floor(u'),$$

where u stands for a user and u’ stands for his/her buddy. Based on different values of a and b, HealthyTogether provides the following three social settings.

- **Competition setting**, where a = 100%, b = 0;
- **Hybrid setting**, where a = 80%, b = 20%;

Figure 4.9: HealthyTogether main interfaces and Fitbit tracker interfaces. Left: HealthyTogether step interface; right: HealthyTogether floor interface.

- **Cooperation setting**, where a = 50%, b = 50%.

In **competition** setting, a dyad competes to gain more points and badges. Thus, a user's Karma points are calculated purely by his/her steps or floors. To gain more badges, both users only need to improve their own activities compared with each other.

In **cooperation** setting, the dyad contributes equally to win points and badges. Thus, we set a = 50% and b =50% so that a user's own performance and that of his/her buddy are equally important for winning badges.

In **hybrid** setting, users both cooperate and compete. We set a = 80% and b = 20%, and thus a user's own performance weights more than that of his/her buddy. We designed a hybrid setting in order to understand the effects of different proportions of competition and cooperation on motivating users.

Badges

HealthyTogether provides incentives by badges. We considered *badges* to be more suitable for two-membered groups than *leaderboard*, which is usually applied in larger groups. The visual design of HealthyTogether badges follows the style of Fitbit badges. These badges are issued based on kp(u). We set a low threshold for the first badge to help users get started easily. For steps, the first badge is issued when $kp_{step}(u) > 500$. This number is followed by 1000 and 2000 and then increases by every 2000 points. For floors, the first badge is issued when $kp_{floor}(u) > 2$, followed by 5 and then increases by every 5 points.

Social Interface

The social interfaces of HealthyTogether v2.0 in three settings are the same except the explanations on how to earn the next badge. Figure 4.10 illustrates an example of step interface. In all three versions, the dyad's current steps (1027 vs. 4388) are visualized as opponents in the first progress bar. However, the second progress bar shows how the dyad's performance contributes to the next badge in different ways. In competition setting, the second progress bar only shows current steps (1027) relative to points for the next badge since the first progress bar already contains necessary information for competition within the dyad; in hybrid badge, the progress bar shows 80% of the user's steps and 20% of the buddy's steps relative to points for the next badge (821 vs. 877); in cooperation setting, the progress bar shows 50% of both the user's and the buddy's steps relative to points for next badge (513 vs. 2194). The area in between of the two progress bars explains how to earn badges. In competition mode, the rule is most straight-forward and thus no additional information is displayed; in hybrid and cooperation settings, this area illustrates the weights of each member's steps, i.e., 80% vs. 20% or 50% vs. 50%. In the example of Figure 4.10, the user needs 973 steps and 377 steps for the third badge in competition and hybrid settings respectively. By contrast, the user has already earned the third badge in cooperation setting.

Users can send messages to each other using HealthyTogether (Figure 4.11a)). They can also choose a template message when they click the template icon located left to the text field. These templates include two themes: cheering and taunting buddies. Figure 4.11b) is an example of template messages for cheering buddies. Users can access messaging interface (Figure 4.11a)) by clicking either cheering or taunting icons, as shown in Figure 4.9. The results of clicking the two icons are the same, but the provided templates are different: either cheering or taunting themes.

4.4.2 User Study

Participants

We recruited participants by posting advertisement on campus (in Switzerland) seeking users who were interested in using Fitbit. After one person signed up, we asked him/her to invite a buddy of his/her choice to join. In this way, we recruited 36 participants, consisting of 18 dyads (two individuals). We required that each dyad should not work in the same office or too close to each other, so that they are less likely to discuss about their experiences face-to-face. We

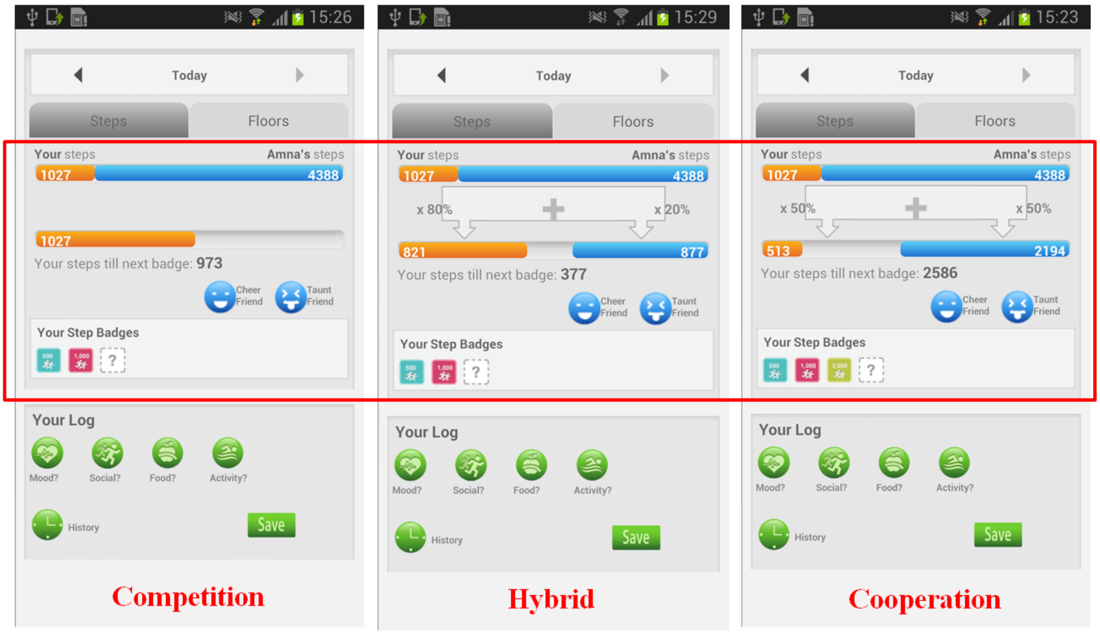


Figure 4.10: Interfaces of the three social settings. Left: competition setting; middle: hybrid setting; right: cooperation setting.

randomly assigned the 18 dyads to one of the three conditions, which led to six in competition setting, six in the hybrid setting and six in the cooperation setting. Hereafter, we refer to participants in the different settings as competition group, hybrid group, and cooperation group respectively.

The participants (15 males and 21 females) are from 17 different countries (Switzerland, France, Iran, China, US, etc.). Most of them are students pursuing different levels of educational degrees (bachelor, master or Ph.D.), except four lab technicians, one researcher and one engineer. All but three age between 20 and 30: one below 20 and two above 30. All of them hold normal body mass indexes (BMI) (18.5-25), except one below normal (15-18.5) and another obese (30-40). They are all currently using a smart phone. No one has used Fitbit trackers before, but six of them have used various mobile fitness apps, such as RunKeeper and Runtastic. Each dyad was made up of close friends. The distribution of gender composition (female-female, female-male, and male-male) was equivalent across dyads in the three social settings (see Table 4.3).

Materials

We provided users with an Android phone with 3G SIM card and a Fitbit tracker, either Fitbit Ultra or its predecessor Fitbit One, but the differences of the two models are irrelevant with our study. Ideally, users could use HealthyTogether on their personal phones. However, as some participants complained either themselves or their buddies did not have an Android

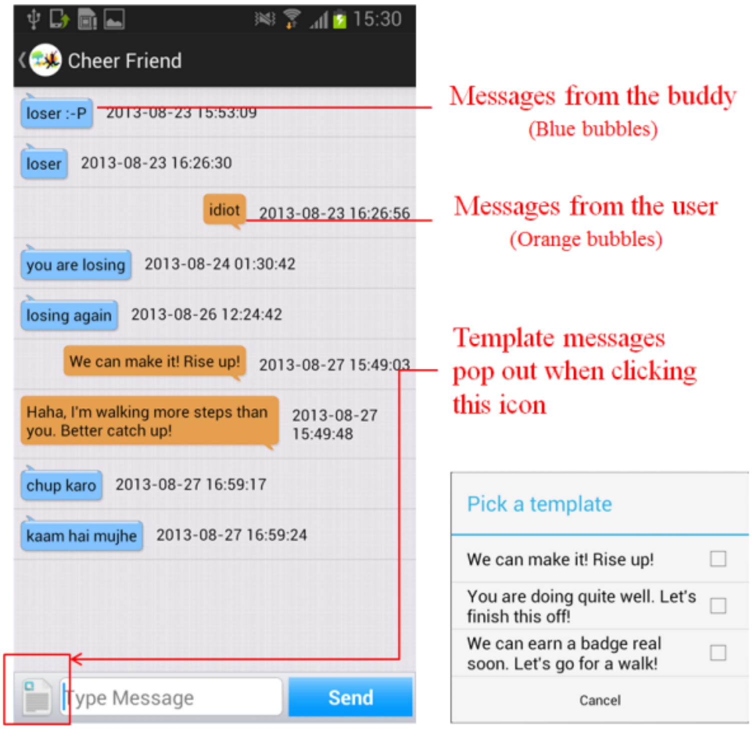


Figure 4.11: Messaging functions in HealthyTogether v2.0. Left: messaging interface; right: templates for cheering buddies.

phone, we decided to provide Android phones to all of them.

Procedure

The study consists of a two-day warm-up session, a one-week control session and a following one-week experimental session. Both control and experiment sessions contain five continuous working days of a week. We refer to the control session as *Phase I* and experiment session as *Phase II*. We chose a period of two weeks so that the control and the experimental sessions span over identical days of the week, thus minimizing the influence of a given day's schedule to the physical activities being monitored. For example, a user may work in the office on Mondays but conduct experiments in the laboratory on Wednesdays. During the warm-up session, participants filled in a demographic questionnaire, signed an informed consent, set up Fitbit account and familiarized themselves with the device. In Phase I, all participants used Fitbit alone without connecting with buddies. Phase I allows us to collect users' baseline activity data. Participants were not informed of HealthyTogether until the end of Phase I, when we visited them individually and distributed the phones installed with HealthyTogether. In Phase II, participants started using HealthyTogether with their buddies in the respective social rules: competition, hybrid or cooperation.

We requested participants to fill in a daily diary in both phases. At the end of each day,

Groups	F-F	F-M	M-M	Total
Competition	3	1	2	6
Hybrid	3	2	1	6
Cooperation	2	3	1	6

Table 4.3: The distribution of dyads with different gender combinations (F-F: female-female, F-M: female-male and M-M: male-male) in three social settings.

we sent a reminder email with the diary link to participants asking them whether they had anything to share with us about their experience using Fitbit or HealthyTogether. The daily diary not only helps us to gain an in-depth understanding of users' experience, but also provides contextual information to participants' activity data during the corresponding day. At the end of the study, participants returned to our laboratory. After collecting back devices, we conducted and audio-recorded an unstructured interview about their overall experiences. We then offered each participant a 40-CHF gift card as compensation for their time.

4.4.3 Findings

In this section, we report results in four aspects: 1) a comparison between physical activities in control session (Phase I) and experiment session (Phase II); 2) a comparison of exercise performances among the three social settings; 3) a comparison of exchanged messages among the three social settings and 4) the relationship between exercise performance and interactive messages. The results were analyzed based on four sources of data: 1) Fitbit activity data, 2) messages exchanged between exercising buddies, 3) daily diary entries and 4) post-study interview recordings.

We prepared a PHP script to Fitbit data of the 36 participants within the two weeks of study. Participants reported two accidents when they forgot to wear Fitbit on a certain day during Phase I, because they changed their trousers. We discarded their step and floor data of the two days from our analysis. The mean daily step count is 9124.3 ($Max = 28816, Min = 614$) and mean floor count is 20.5 ($Max = 85, Min = 0$). We coded each entry with the phases, i.e., Phase I or Phase II. We also coded them with users' social setting, i.e., competition, hybrid or cooperation.

Additionally, we gathered a total of 370 messages that participants sent when using Healthy-Together in the five working days of Phase II. Most of the users sent messages in English except eight in French, two in Urdu and two in Persian. We recruited volunteers with the above as native language to translate the messages into English. Afterwards, two researchers inductively coded the messages.

Finally, we deductively analyzed daily diaries and post-study interviews, based on findings from quantitative analysis. The length of diary entries ranges from a short sentence (e.g., 2 words) to an essay (e.g., 1272 words). The length of the post-study interview lasted up to 20 minutes. The audio-recording was transcribed by the researchers. Qualitative data, such as the diary entries and post-study interview, not only provide complementary and contextual information to quantitative data, but also help us understand why and how our system and social settings work [80].

Individual versus Team Exercise

We first compared participants' exercise in control session and experiment session. Results of a paired-samples t-test showed that users' steps have significantly increased by 11.8% from Phase I (M = 8828) to Phase II (M = 10340), $t(36) = 5.12, p < .001$. Similarly, the average floor count has significantly risen by 15.1% from Phase I (M = 20.00) to Phase II (M = 23.03), $t(35) = -3.22, p = .003$. If we assume that Phase I is a baseline for participants, then the above results indicate that **HealthyTogether helps users to increase their activity levels in terms of steps and floors**.

Daily diary also reflects the above findings. When users started to use Fitbit, they were highly motivated, because Fitbit helped them to quantify themselves and increase their awareness of physical exercises.

> "*The device is cool, and it did increase my incentive to walk around. I didn't know that I can make so many steps in a single day!*" – P12

However, some issues arose when they used Fitbit individually. Some reported that they forgot paying attention to Fitbit when they were busy, since Fitbit was unobtrusive to wear.

> "*Unfortunately today was such a busy day that I didn't even have time to think about Fitbit.*" – P25

Several participants also mentioned that the novelty effect of Fitbit wore off after a few days.

> "*Not as great as the first two days.*" – P9

Meanwhile, a few advantages emerged when they were using HealthyTogether. First, HealthyTogether has helped them to compare with each other. Participants widely appreciated that they were motivated to pay attention to their physical activities because someone else was present.

> "*From today on, my buddy is watching me. I was motivated to spend the day with more vigor and less laziness.*" – P34

Additionally, participants were triggered to compare with their buddies.

> "*I have been frequently checking and comparing with how she was doing.*" – P7

Furthermore, they could send messages to each other.

> "*She made me aware that it was coffee time and I should take a break, go upstairs and take a coffee*" – P33

The above findings confirm the qualitative results from Consolvo et al. [42] and Toscos et al. [113] that applications that integrate social influence and message exchange can motivate users compared when they exercise alone.

Groups	Data Type	Phase I (Individual)	Phase II (Team)	% Inc.	p-value
Overall	Step	9332	10433	**11.8**	***.002***
	Floor	20.00	23.03	**15.1**	***.003***
Competition	Step	9687	10181	5.1	.507
	Floor	19.74	21.48	8.8	.300
Hybrid	Step	10064	11148	**10.8**	***.061***
	Floor	19.05	22.51	**18.2**	***.066***
Cooperation	Step	8243	9971	**21.1**	***.001***
	Floor	21.22	25.10	**18.3**	***.039***

Table 4.4: A comparison of different groups: average steps and floors in Phase I and II, percentage increase from Phase I to II, and p-value of 2-tailed paired-samples t-test between Phase I and II.

Competition versus Cooperation

We then examined the impact of the three social settings in HealthyTogether. A paired-samples t-test was conducted to compare mean step and floor count in Phase I and Phase II in three groups respectively (see Table 2). There was a significant step increase in both hybrid groups ($\Delta = 1084$, $t(11) = -2.09$, $p = .061$) and cooperation groups ($\Delta = 1728$, $t(9) = -4.29$, $p = .001$). The floor increase also tends to be significant in hybrid groups ($\Delta = 3.5$, $t(11) = -2.04$, $p = .066$) and cooperation groups ($\Delta = 3.9$, $t(9) = -2.34$, $p = .039$). However, no significance was found when comparing the two phases in competition groups. This suggests that **both cooperation and hybrid setting could effectively motivate users to do more physical activities**, i. e., taking more steps and climbing more stairs.

We further compared whether the effects of cooperation and hybrid settings were significantly different. We were aware that participants had diverse lifestyles which could lead to different levels of steps and floors between individuals. Thus, we used percentage increase from Phase I to Phase II as the metric for comparative analysis. Results of an independent-samples t-test showed no significant difference in percentage increase between cooperation and hybrid groups (21.1% $vs.$10.8%, $t(22) = -1.46, p = .159$). Nor did we find significance for floor percentage increase between the two settings, (18.3% vs. 18.2%, $t(22) = .095, p = .925$).

Additionally, we analyzed the performance between two participants in a dyad. The percentage increase of users' steps did not correlate significantly with that of their buddies, and nor did floors. Interestingly, in competition groups, percentage increase of steps between teammates tended to be negatively correlated, $r = -.908, p = .012$, and so did floor increase, $r = -.57, p = .085$. This indicates **a reverse tendency of step enhancement between teammates in competition setting**: the more one increase, the less likely his/her buddy would do.

Analysis from users' diaries and post-study interviews showed evidence of the above findings. Generally, participants in **cooperation** group reported positively about the experience of exercising with their buddies. Firstly, they not only focused on their own performance, but also attempted to encourage each other. For example, P34 reported,

> "*Usually I prefer to stay at my office to work. But because I knew I'm playing with someone, I went out with my papers going to the farthest seats in the building and studied there. Moving around a little bit refreshed my mind. So I shared my experience with my buddy hoping she could walk more.*"

Additionally, participants felt obliged to increase their activity levels when their performance was connected with their buddies. As is written by P33,

> "*I realize that i am always winning the same badge as my buddy! We live almost same distances from the university. Seeing her always having more steps makes me a bit guilty. Maybe i should take more breaks and move around.*"

In addition to helping each other, participants also tended to be comparable with their buddies in **hybrid** group. As is reported in P24's diary,

> "*I ran after work because during the day i had course from morning till afternoon so that i was not really moving. When my partner saw that i achieved 19000 steps he also went out to run...he was jealous.*"

Several participants in **competition** group also enjoyed the experience with their buddies. For example, P9 lived in the same community as his buddy (P10), but always lost the competition. So he had analyzed the causes:

> "*I think I'll lose the competition against my buddy for some ergonomic and physical reasons: she is shorter than me; hence, with smaller legs. Of course, she needs more steps to get the same places as me. Therewith, by considering a similar daily routine of 'in-out', I'm very prone to lose.*"

Despite of this, he had been industriously boosting his climbing performance.

> "*I was really motivated to increase my stairs and really committed to strive some success, and I made it*".

His buddy also reported enjoying winning the competition and exchanging messages with each other.

Design Principle 4.4.1. *Consider designing cooperation in gamified exercise applications.*

As is consistent with quantitative results, some participants reported some demotivating factors of **competition**. For example, P2, an active person who vowed to win the competition at the beginning, stated in post-study interview,

> "*I was very enthusiastic to compete with Ken. On the morning of the first day, I found his step more than mine. No way! I will beat him! But later it was not challenging at all. He was always sitting in front of his computer. I want to compete with some students who are active in the sport center and have regular exercises.*"

On the other hand, his buddy (P1) described:

> "*He was so competitive and energetic. I don't think I can ever compete with him.*"

Similarly, P11 wrote in her diary:

> "*I expected having a buddy to encourage me to be more active. It didn't happen so far :(*"

Her teammate P12 reflected in her post-study interview,

> "*Competition is not helpful if one day your buddy has no time for exercise. Competition is good only if you have competitors who are more or less in the same situation as you. If I have a friend who runs 20 Km per day, I'm not interested in competing with him/her.*"

Similar stories were reported by seven participants among the twelve in competition setting. Competition groups' experience shows that **competition motivates dyads if they have equivalent performances and availabilities, but is likely to demotivate them if otherwise**.

Design Principle 4.4.2. *When designing playful competition, consider pairing partners with equivalent abilities.*

Messages

Participants sent a total of 370 messages when using HealthyTogether in five working days, shared by competition groups ($n = 58$), hybrid groups ($n = 114$) and cooperation groups ($n = 198$). Their average number of messages in different social settings is listed in Table 4.5. A one-way ANOVA test showed significant difference among the three settings, $F(2,33) = 7.17, p = .03$. Post hoc (LSD) test revealed that cooperation groups sent significantly more messages than competition groups ($p = .001$) and hybrid groups ($p = .030$). More concretely, the number of messages in cooperation groups is 73.6% more than hybrid groups and twice more than competition groups. However, no significance was found between hybrid and competition groups. This indicates that **cooperation setting is more likely to stimulate users to interact with each other via messages than competition and hybrid settings**.

Groups	Mean	SD	Max	Min
Overall	10.50	9.22	36	0
Competition	4.83	5.20	19	0
Hybrid	9.67	7.81	26	2
Cooperation	17.00	10.02	36	3

Table 4.5: Descriptive statistics of messages.

We then examined the content of the messages that users have exchanged. Table 4 shows six themes that emerged. Cheering buddies had the largest share ($n = 126$, $pct = 34.1\%$), which was three times the number of taunting messages ($n = 42$, $pct = 11.4\%$). Noticeably, users also shared their experience in physical exercises ($n = 75$, $pct = 20.3$). This included how they have successfully gained more steps, what caused their step count to drop and how they should improve in the next day.

We also compared the distribution of messages within the six message themes by the three social groups. The distribution, as shown in Figure 4.12, reveals the following phenomenon: 1) for the theme of cheering buddies and sharing experience, cooperation groups have obvious larger proportions, followed by hybrid groups, while competition groups hardly shared experiences with their buddies; 2) users in all the groups chatted and compared with their buddies; 3) competition groups sent more taunting messages than cooperation and hybrid groups; and 4) competition groups hardly organized activities together via messages. The above analysis showed that the three social incentives were likely to drive users to interact with their buddies in different ways.

We further examined the messages about steps and floors among all messages except the theme about chatting. Messages about general activities, steps and floors took shares of 42%, 48% and 10% respectively, indicating that users sent more messages about steps than floors. This might be due to the fact it is more convenient to encourage users to take a few steps than climbing the stairs. For example, one can easily stand up from his/her chair and walk a few

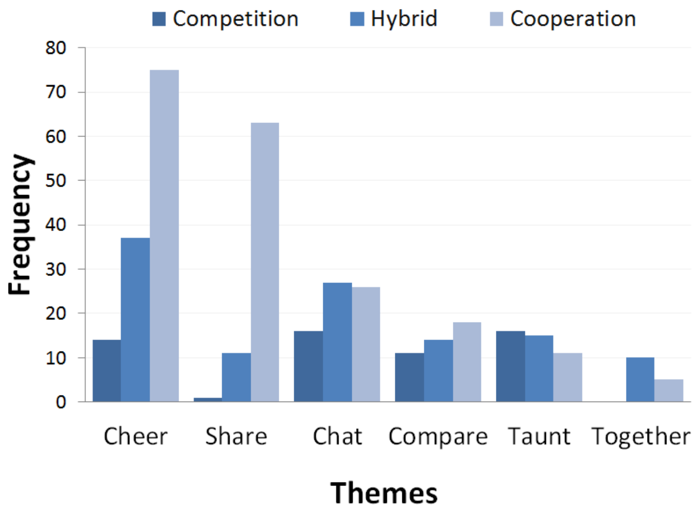


Figure 4.12: Message themes and distributions by group.

steps around the corridor. By contrast, s/he might not have the opportunity to immediately find stairs nearby. However, this surmise needs further validation.

Themes	Examples	%
Cheer	*"u r doing great!"*	34.1
Share experience	*"2k steps more by walking home instead of bus."*	20.3
Chat	*"Good morning!"*	18.6
Compare	*"You are still better than me :("*	11.6
Taunt	*"Loser!"*	11.4
Workout together	*"Coffee together?"*	4.1

Table 4.6: Message themes sorted by proportion.

Design Principle 4.4.3. *Consider providing templates for messaging and tailoring the content for the respective goals of the strategies.*

In addition to themes of messages, we also observed several interesting patterns. Firstly, users frequently used **emoticons**, such as ':o)', ':P'. Messages with emoticons constitute 24.3% (n = 90) of all messages. Secondly, template messages were mainly used to help users get started with communication. Template messages, which constituted 7.6% (n = 28) of the total number, were primarily sent out during the early rounds of message exchange. This reflects several suggestions that participants proposed in order to facilitate them in sending messages. For example, P9 suggested including emoticons in sending messages.

> *"I like the taunting face (icon). If would be nice to use it in the messages. Or perhaps sending a taunting face is convenient."*

For another example, P23 favored templates especially when she did not know how to encourage her buddy.

> *"I don't know what to say, but the template messages helped me to get started."*

This indicates that **designing user-friendly messaging interface could help users to communicate more actively with their buddies**.

Design Principle 4.4.4. *Consider providing emoticons in messaging components to encourage users to express their feelings while playing exercise games with their buddies.*

Physical Activities and Communication

Communication between teammates has been identified as beneficial for enhancing activities by users' *qualitative* feedback [29, 41, 96]. This subsection presents statistical analysis between users' activities and communication. We used **percentage increase of activities** from Phase I to Phase II to measure a user's improvement in physical activities and **message count** to measure his/her communication volume. More concretely, percentage increase includes two facets: steps and floors. For message count, we refer to the messages a user sent as sent messages and messages that the buddy sent as received messages. We conducted a bivariate analysis to examine the association between percentage increase and message count. Table 5 shows the Pearson correlation matrix of the variables.

We first examined the sub-metrics that describe percentage increase and message count. Step percentage increase is significantly correlated with that of floors ($r = .405, n = 36, p = .032$), indicating that users' improvements in steps and floors associated with each other. Users' number of sent messages and that of received messages were also significantly related, $r = .877, n = 36, p < .001$. This hints that the level of a user's social interaction (in terms of communication) is associated with that of his/her buddy.

We then investigated the relationship between users' number of messages and improvement in physical activities. Percentage increase of steps was positively correlated with number of sent messages ($r = .398, n = 36, p = .036$, two-tailed) and number of received messages ($r = .394, p = .038$, two-tailed). Thus, the more a user received and sent messages to his/her exercise buddy, the more likely he was motivated to take more steps compared to his baseline performance; similarly, the more progress s/he made, the more likely s/he was motivated to send messages to his/her buddy. However, we did not find strong correlation between percentage increase in floors and the number of message. This could be due to the fact that messages about steps outnumbered those about floors, as presented in previous subsection. Therefore, **exchanging messages help users to increase their steps**.

As is consistent with quantitative results, users' feedback confirmed that sending messages helped users to become more active in all three social settings. For example, P4 in competition group wrote,

> "*Every day she won, she would send me one message: loser. Finally, I beat her during the weekend because of my hiking trip.*"

P20 in hybrid group also logged,

> "*I told my buddy that I ran along the lake today. It motivated him to go to the gym.*"

> "*Today, I was more eager to take more steps and don't feel guilty anymore! A few messages were enough to persuade me!*" – P34 in cooperation group

Thus, it is essential to stimulate users to communicate with their buddies.

	1	**2**	**3**	**4**
1. Step % increase	1			
2. Floor % increase	.566**	1		
3. Message sent	.339*	.084	1	
4. Message received	.335*	.080	.876**	1

Table 4.7: Correlation Matrix among percentage increase in physical activities and message count ($*p < .05, **p < .01$).

4.4.4 Summary

In this section, we have developed a mobile application called HealthyTogether that allows dyads to participate in daily physical exercises as a game. We conducted an in-situ user study with 18 dyads, over a period of two weeks and compared participants using HealthyTogether in three social settings. The two-week study was divided into a one-week control session and a one-week experiment session using HealthyTogether. Results show that users' activities in experiment session have increased significantly compared to control session. Among the three social settings, cooperation and hybrid outperformed competition in motivating users to enhance their physical activities. More specifically, users have significant step and floor increase in both cooperation (by up to 21.1%) and hybrid setting (by up to 18.2%), but not in competition setting (by up to 8.8%). In addition, increase in physical activities significantly associates with the number of exchanged messages, and users in cooperation groups sent nearly twice the number of those exchanged in hybrid groups and three times more in competition groups. Based on the findings, we summarize four design implications from our studies.

1. Consider designing cooperation in gamified exercise applications.

2. When designing playful competition, consider pairing partners with equivalent abilities.
3. Consider providing templates for messaging and tailoring the content for the respective goals of the strategies.
4. Consider providing emoticons in messaging components to encourage users to express their feelings while playing exercise games with their buddies.

This work can provide researchers with an environmental platform for testing social schemes and serve as a starting point to investigate social incentives. It can also help practitioners to design persuasive technologies that enhance social user experience. In the future, we will examine more diverse social incentives and investigate group dynamics in larger groups.

4.5 Social Capital as Social Incentives

Current pervasive fitness applications employ various strategies, such as virtual badges, goal-setting and community competition, to motivate users [80]. However, most of these strategies are incentives for individual effort. In this experiment, we explore the effectiveness of social capital as incentives for behavior change. Social capital refers to "the resources that are created through interactions in social relationships and that provide benefits to participants in the network." [40] In particular, we are interested in the following research questions:

- Are users more motivated to exercise by winning badges for a team or themselves?
- Would users behave differently when exercise with a buddy for a shared goal from for an individual goal?
- Would users engage more in community competition based on teams or based on individuals?

To understand these questions, we integrated interfaces and incentives that involve social capital to HealthyTogether as an experimental platform. Furthermore, we are interested in evaluating the intervention effect over a longer period compared with studies in Section 4.3 and 4.4. To this end, we conducted a controlled study over a period of ten weeks with 20 users. In this section, we report the findings of the user study and derive design guidelines. In the rest of the section, HealthyTogether stands for the version 3.0.

4.5.1 Experimental Platform: HealthyTogether v3.0

We introduced three new features to HealthyTogether v3.0: team badges, shared goal-setting, and team leaderboard. The rest of the functions remain the same as HealthyTogether v2.0 (see Section 4.4). The main interfaces of HealthyTogether v3.0 are shown in Figure 4.13.

Figure 4.13: HealthyTogether interfaces. Left: team page; right: team leaderboard.

Team Badges

The first feature – team badges – allows dyads to win virtual badges collectively. We adopt the "cooperation" rule that is demonstrated to be most effective in the studies in Section 4.4. A user can win step and floor badges based on his/her Karma points, which are calculated as following.

$$kp_{step}(u) = 50\% \times step(u) + 50\% \times step(u')$$

$$kp_{floor}(u) = 50\% \times floor(u) + 50\% \times floor(u'),$$

where u stands for a user and u' for his/her buddy. In this setting, the dyad contributes equally to win points and badges. Thus, we set a = 50% and b =50% so that a user's own performance and that of his/her buddy are equally important for winning badges.

Same as HealthyTogether v2.0, these badges are issued based on $kp(u)$. For steps, the first badge is issued when $kp_{step}(u) > 500$. This number is followed by 1000 and 2000 and then increases by every 2000 points. For floors, the first badge is issued when $kp_{floor}(u) > 2$, followed by 5 and then increases by every 5 points.

Collective Goal-setting

Collective goal-setting allows users to set accumulative daily exercise goals for the two as a dyad. Users can pledge the number of steps and floors by choosing from a list of options. Once one user of the dyad sets the goal, it cannot change for the day.

Figure 4.13 (left) displays the social interface for collective goal-setting. A progress bar visualizes each user's proportion of his/her steps or floors in achieving their goals. The left part represents the user's proportion in fulfilling the team-goal and the right part shows his/her buddy's contribution in the goal. For example, in the case of Figure 4.13, the user has 9791 steps while his/her buddy has 6659 steps. As a team, they completed 16450 steps, which is 82% of the number of steps they have pledged for the day (20000). Since the goal is set for the sum of two users' number of steps or floors of a day. Thus, we call it collective goal-setting.

Team Leaderboard

The team leaderboard enables dyads to compete with other dyads in a community. This is different from the conventional leaderboard design, in which users are ranked in the list by their own performance and statistics. This feature is inspired both by users' suggestions in Section 4.3 and by the Olympic Games, which not only provide individual games, but also team events.

Users can access the team leaderboard by clicking the yellow community button in the main page of HealthyTogether (see Figure 4.13 (right)). They can set an avatar to represent their team. The leaderboard page displays the sum of steps or floors of all HealthyTogether dyads in a ranked list. The progress of the user's team is visualized in orange while that of the other teams is in blue. When users click the progress bar, they can check the name of the team. In Figure 4.13 (right), the user's team (visualized in orange progress bar) ranked the second place in the leaderboard for steps, with a total number of 16450 steps. The team at the top of the leaderboard has walked a sum of 19745 steps.

4.5.2 User Study

Study Design

We conduct a longitudinal controlled study to evaluate the effectiveness of social capital in pervasive fitness applications. We deploy the study using single-case design, which are commonly used for testing experimental technologies for behavior change. In single-case design, each participant serves as his or her own control [47]. The controlled phase and intervention phase are repeatedly deployed to evaluate the effectiveness of the intervention.

In this study, we refer to the period when participants using Fitbit sensor and dashboard as controlled phase, and the one when they use HealthyTogether as experiment phase. Section

4.5.1 extensively describes incentives such as virtual badges, goal-setting, and leaderboard at the *team* level. As a control condition, we employ Fitbit application to study the above incentives at the *individual* level. Fitbit has a dashboard that not only displays users' activity data, but issues badges, supports goal-setting, and provides a leaderboard that ranks users' sum of steps in the recent seven days. The dashboard is available for both web version and mobile application.

Participants

We recruited participants by posting fliers on campus (in Switzerland) seeking for users who are interested in using Fitbit. After four individuals contacted us and registered in the study, they introduced their friends via word-of-mouth. Therefore, many of the participants loosely know each other to different degrees. In this way, we recruited 20 participants, including nineteen students (both Bachelor and Master) and a researcher. Their ages range between 19 and 40, with eleven males and nine females. They came from seven different countries (Switzerland, France, Macedonia, Turkey, etc.). All but one user currently use a smartphone. None of them have used Fitbit before, but six of them used fitness related devices and applications (e.g., RunKeeper, Wii), and fourteen of them have heard about such tools.

Materials

We provided users with an Android phone and a Fitbit tracker, either Fitbit Ultra or its predecessor Fitbit One, but the differences of the two models are irrelevant with our study. Since the study lasts for ten weeks, some participants (N=9) who have an Android phone prefer to use their own phones.

Procedure

The ten-week user study is divided into three phases. In the controlled phase (Phase I), which lasts for four weeks, participants used Fitbit only. In the experiment phase (Phase II), a period of four weeks, they paired up with a buddy to use HealthyTogether as a dyad. This is followed by a second control phase (Phase III), when they stopped using HealthyTogether and only used Fitbit for two weeks.

After the participants registered in the user study, we invited each of them for a 30-minute initiation meeting in our lab. During the meeting, we first debriefed the goals and terms of condition of the user study. Upon signing the informed consent form, we conducted a semi-structured interview to survey their background, including their motivation and experience in using pervasive fitness applications. At the end of the meeting, we provided each participant with a Fitbit tracker and briefly showed them its functionalities and installation.

Towards the end of one month's usage, we contacted each participant to schedule the second meeting. We also introduced them HealthyTogether and encouraged them to pair up for the second month of the study (Phase II). During the second meeting, we interviewed them about their experience with Fitbit during the first month. The interview lasted for around 30 minutes. Afterwards, we provided with an Android phone (unless they requested to use their personal phones), and helped them download and install HealthyTogether from Google Play. When they registered their accounts in HealthyTogether, they sent out or accepted an invitation to pair up with their buddy via HealthyTogether. All participants started HealthyTogether in the same week to facilitate the team competition in the community. Except for one dyad, each user knows his/her buddy to different degrees.

At the end of Phase II, we invited them to our lab for the third meeting. We started with a 30-minute interview about their experience with HealthyTogether and asked them to compare Fitbit application and HealthyTogether. After the third interview, we requested them to use Fitbit application – without using HealthyTogether – for another two weeks.

When they finished the study after around ten weeks, we gave Fitbit or a 50-CHF gift card to the participants as compensation for their time, up to their choice. In the end, all but four participants chose to keep Fitbit as their personal belongings. All the three interviews were audio-recorded and transcribed for further analysis.

4.5.3 Findings

In this section, we report results in two aspects: 1) a comparison between physical activities in the control sessions (Phase I&III) and the experiment session (Phase II) and 2) users' qualitative evaluation and comparison between individual and social incentives for the three strategies. The results were analyzed based on Fitbit activity data and interview recordings.

We prepared a PHP script to Fitbit data of the 20 participants within the ten weeks of study. Three users were not able to join Phase III due to personal plans, such as long summer vacation and losing Fitbit. The mean daily step count is 8576 (Max = 38043, Min = 614) and mean floor count is 20.4 (Max = 554, Min = 0). We coded each entry with the phases, i.e., Phase I, Phase II and Phase III.

We then transcribed and deductively analyzed three interviews to identify differences (if any) and the reasons. The total length of the interview recordings is up to 26 hours. Qualitative data gathered from empirical studies of pervasive health applications can not only provide contextual explanation to quantitative data, but also help us understand why and how social strategies work [80].

Intervention Effect

We first compare users' number of steps and floors throughout the study and in the first two phases. Overall, participants' steps increased significantly from Phase I to Phase II ($M1_{step}$ =

$8453, M2_{step} = 8996, N = 20, p = .08)$. Their number of floors increased from $(M1_{floor} = 20.5)$ in Phase I to $(M2_{floor} = 21.7)$ in Phase II, but the increase is not statistically significant. This confirms with our findings in Section 4.4 that **HealthyTogether helps users to do more physical activities compared to when they use Fitbit application**.

We then compare users' steps and floors between repeated baselines (Phase I&III) and intervention (Phase II) periods. We analyze the data of 16 users with complete Fitbit data in all three phases. The other four users had unexpected incidents and thus we were not able to collect their complete data in the Phase III. Two users left for a vacation, one user forgot to synchronize data and then lost his Fitbit and the other user changed to another version of Fitbit which does not track the number of *floors* in Phase III. Table 4.8 shows the average number of steps and floors in the three phases and the increase between phases. Users' steps have increased significantly by 9.5% from Phase I to Phase II $(M1_{step} = 8281, M2_{step} = 9069, \Delta = 788, N = 16, p = .026)$, and their floors slightly increased by 3.3% without statistical significance $(M1_{floor} = 18.3, M2_{floor} = 18.9, \Delta = 0.6, N = 16, ns)$. Align with this results, their average number of steps decreased significantly by 20.8% from Phase II to Phase III $(M2_{step} = 9069, M3_{step} = 7184, \Delta = -1885, N = 16, p = .00)$, and number of floors decreased by 15.3% $(M2_{floor} = 18.9, M3_{floor} = 16.0, \Delta = -2.9, N = 16, p = .007)$. The results of repeated baseline study demonstrated **the effectiveness of HealthyTogether in engaging users in physical exercises *in the long run*.** The intervention effect is more obvious for walking (steps) compared with climbing (floors), maybe due to the fact that HealthyTogether interface shows steps in the first tab.

	Overall	Phase I	Phase II	Phase III	Δ(I→II)	Δ(II→III)
Steps	8371	8281	9069	7184	788*	-1885**
Floors	18.0	18.3	18.9	16.0	0.6	-2.9**

Table 4.8: Average number of steps and floors per day overall and in the three phases ($N = 16$, $^*p < .05$, $^{**}p < .01$).

We then examine each individual's physical activities ($N = 20$) across the three phases. Figure 4.14 and 4.15 visualize each users' average number of steps and floors in the three phases. All but five users have walked more in experiment stage (Phase II) compared with the first baseline (Phase II). Among them, six participants have their number of steps increased by an average of more than 2,000 per day and two have increased by an average of more than 1,000 per day. For controlled evaluation of intervention to promote physical activities, an average daily increase of 1000 is considered significant [75]. For floors, all but seven have increased their average number from Phase I to Phase II. Among them, two users have increased an average of more than 15 floors per day. Based on users' feedback during the third interview, most of them reported that they were busy preparing for the exams during Phase II and had less chance to allocate time for physical activities. Despite that, we still observed the obvious improvement. Based on the qualitative results, we derive the following design principle:

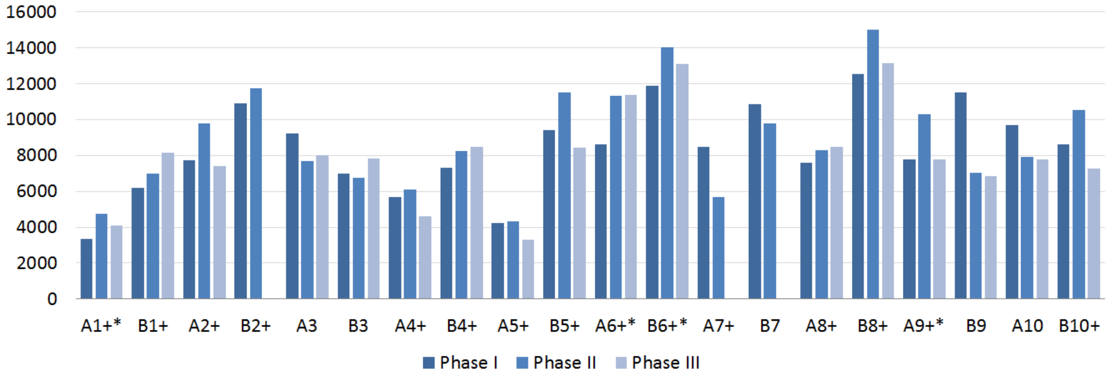


Figure 4.14: Distribution of average steps per day in three phases by dyads. +:steps increased from Phase I to Phase II (*$p < .05$).

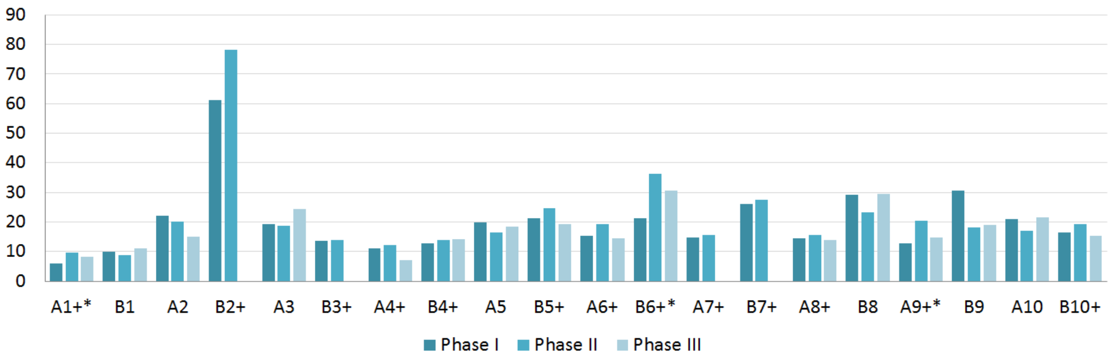


Figure 4.15: Distribution of average floors in three phases by dyads. +:floors increased from Phase I to Phase II (*$p < .05$).

Design Principle 4.5.1. *Consider designing incentives that involve social capital to engage users in the long term.*

Design Implications

After discovering the positive intervention effect, we are interested in identifying the factors that cause the change. We report findings based on users' qualitative feedback collected during the interview.

Overall, users provided positive feedback for HealthyTogether. They usually check HealthyTogether on a regular basis, e.g., in the morning, at noon and in the evening. But they are more motivated to use HealthyTogether when they finish some activities, such as *"when I walked a lot"* (B1) or *"when I finished hiking"* (B2). Some users like the fact that HealthyTogether is inherently designed *social* and helps the dyads to know about each other's schedule and patterns.

> *"In HT, you could see what's your partner is doing, you can see the others' amount. Also in HT, you can see everybody. In FT, you only see people you add. It's easier to taunt them in HT. In HT, you did not have to add everybody, you can see everybody without knowing them. You want to be the best one in the app, more challenging than your friends. More challenging on HT than Fitbit."* – A3.

We then compare Fitbit application and HealthyTogether in the three aspects: badges, goal-setting and leaderboard.

Badges When comparing individual badges (in Fitbit) and team badges (in HealthyTogether), users did not report much difference in motivating them. As Badges bring them more fun and playfulness. As B1 said,

> *"It was nice to earn badges. But I was not specifically walking to earn badges. Badges are not the main motivation; they just make me happy when I get them. it is not a goal."*

Despite little perceived differences, four participants mentioned that one merit of Fitbit is that it also provides badges for accumulated steps or floors. For example, Fitbit issues "lifetime badges" whenever users' total steps or floors reach a new threshold. This can motivate users in the long run compared with issuing badges on a daily basis.

> *"In Fitbit, they have calculated all the steps you have done. I had a badge for that many stairs; for HealthyTogether, the counting of badges is reset every day. That's motivating as well on that day. If you are using both, it's indeed a long term commitment."* – A3.

The above finding leads to the following design principle.

Design Principle 4.5.2. *Consider designing badges to motivate and sustain users' exercising performance for both short-term (e.g., a day) and longer period (e.g., one week or 'the lifetime') in social gamification.*

Goal-setting Most participants (N = 11) preferred the shared goals in HealthyTogether than individual goals in Fitbit for various reasons. Some like the sense of **teamwork and accountability** in collective goals.

> *"I'm happy when I finish it, for both. For Fitbit, you are alone and not motivating. Nobody is helping you."* – B9.

> *"Team setting is more interesting, maybe because I feel we have a goal together, and I don't want to fail for a team. It's like having responsibility for one another."* – B1.

Some users also revealed that they became more **familiar** with buddy's overall daily physical activities in order to make a reasonable pledge.

> *"We used it based on what we did the day before. Sometimes I set up a default goal or something that you are successful in your goal before."* – A1.

Some users ($N = 4$) complained some difficulty to make a pledge for both of them especially if they were not very familiar with each other. For example, P4, who did not know his buddy before, commented:

> *"Some randomness here. You don't know whether you can do it or not. If you could both pledge something, and then the goal is chosen. The days that I could pledge is a bit random. Let alone for my buddy."*

Thus, some suggested setting the collective goal by asking both members to set individual goals and making the sum of them.

> *"It would be nice if there are two goals. You want to do some sport, but your buddy is busy. If both have an individual goal and then collective goal."* – A9.

Alternatively, it can be more attractive if the teams can challenge other teams for the collective goal.

> *"Maybe we could compete with the other groups. If I challenge someone, we can challenge the others, we can go against the other groups. You can have a goal. The other groups make 6000 today, or we can challenge then that we will make more."* – A7.

Users' feedback has shed light on many promising strategies that could make collective goal-setting more fun. Thus, we suggest:

Design Principle 4.5.3. *Consider designing collective goal-setting to increase the perceived playfulness and social accountability; when implementing the design, consider summing up all group members' individual goals to make the goal more achievable and controllable.*

Leaderboard The team leaderboard is reported to be the most desirable feature in HealthyTogether. The major reason is reported to being able to cooperate with another one and compete in a community, regardless whether the buddy knows each other or not.

> *"It's fun to compare and with other groups. Usually we are first. I did most of the steps, and my buddy did most of the stairs – he had average 20-25 floors per day. For steps, usually I have more. It's good to see we are always first. I tried to look where we were in the Fitbit leaderboard, but the leaderboard with buddy is much more fun, especially with people you know. Even without communication, it is interesting enough. But in case I know my buddy, I would interact with the him, maybe using HealthyTogether or Facebook."* – A2 (who does not know his buddy before).

> *"That's really funny. It is the most motivation to see other teams. We usually ranked 2nd or 3rd. sometimes at the top. It is good to see that your team is better than other people. "* – B9 (who knows her buddy).

Another frequently mentioned reason is the *fun* part to collaborate with another one to compete in a larger community.

> *"The team leaderboard is more important than the individual one, I don't know why. I did not pay attention to the individual ranking. Maybe I did not have many friends, I cannot compare with many people. Even though I don't know them, it's not a problem. Fitbit is more important to see myself where I stand. The team leaderboard is more fun to see. It's the fun part."* – B1.

> *"I like HealthyTogether is that you can see every day where you are. On Fitbit, if there is a person did 15k per day, he will be there the whole week. It's fun to have group competition, we can say 'we did it' and fun to cheer each other. Sometimes, I asked my buddy to walk a bit to surpass someone."* – B4.

While some users ($N = 5$) acknowledged that implicit and 'silent' community competition is already very exciting, some ($N = 3$) suggested providing communication features to allow them to interact with other groups, such as sending messages, poking them and challenging them.

> *"It might be interesting to challenge the other team. For example, you go to community, you click on the team name. I challenge you to do 20k steps. The other team to walk together to beat you."* – B5.

> *"It's nice, but I would prefer to have a connection with them. Just to see an icon of them. Make me feel somebody else is out there."* – A10.

Some participants ($N = 2$) also provide design alternatives to visualize leaderboard instead of displaying them all in a list. One participant (A7) sketched a visualization method that highlights those winning teams in an "rewarding stage" with champions, runner-ups and third place, and listing all the rest teams to create more fun.

While participants who expressed the motivating factors in the interview also improved their steps in quantitative analysis. We also observed the significant decrease of steps for participants who showed less motivation from "teamwork". For example,

> *"We have less challenge when we are in a group because if I don't make any steps, maybe my buddy will have more. So it's more motivating to do it alone. In group, it is not only your fault. If you don't have a lot of steps, it might be my buddy who is not walking."* – B6.

Based on the evident positive feedback we received from the users, we suggest:

Design Principle 4.5.4. *When designing leaderboard, consider team competition and communication in the community to enable group dynamics.*

4.5.4 Summary

In this work, we compare deploying strategies for pervasive fitness applications for individuals and for social capital. The strategies include goal-setting, virtual rewards (i.e., badges), and community competition (i.e., leaderboard). We employed Fitbit application and a mobile social game called HealthyTogether as experimental platforms, both of which aim to motivate users to walk more and climb more stairs. We conducted a longitudinal study consisting of a four-week individual controlled phase, a four-week social intervention phase, and another two-week individual controlled phase involving 20 participants. Results show that users have significantly higher number of daily steps in the intervention phase ($M2_{step} = 9069$) than that in the controlled phases ($M1_{step} = 8281, M3_{step} = 7184$). Results also show a slightly higher number of daily floors in intervention phase ($M2_{floor} = 18.9$), compared with controlled phases ($M1_{floor} = 18.3, M3_{floor} = 16.0$), but the difference is not statistically significant. Users' qualitative feedback indicates an overall preference towards incentives in social setting than those in individual setting. More concretely, they found competing with other teams in the community is more playful and stimulates more interactions with the buddies. Additionally, setting a goal with a buddy brings more accountability, familiarity towards each other. For badges, they find it encouraging when receiving them but not the principal motivation. Based on the findings, we derive four design principles for persuasive fitness applications.

1. Consider designing incentives that involve social capital to engage users in a long term.

2. Consider designing badges to motivate and sustain users' exercising performance for both short-term (e.g., a day) and longer period (e.g., one week or lifetime) in social gamification.

3. Consider designing collective goal-setting to increase the perceived playfulness and social accountability; when implementing the design, consider summing up all group members' individual goals to make the goal more achievable and controllable.

4. When designing leaderboard, consider team competition and communication in the community to enable inter-group dynamics.

4.6 Chapter Summary

In this chapter, we study the design of social interactions in persuasive technologies for physical activities. We refer to the social interaction mechanisms that are designed to stimulate users' physical activities as social incentives. To examine and compare various social incentives, we developed a mobile fitness application called *HealthyTogether* that allows dyads to walk together. We conducted three empirical studies with 80 users with a period of up to ten weeks. Our study results of the three studies unanimously demonstrate of the superior effectiveness of deploying social incentives compared with individual settings. After studying social incentives such as accountability, competition, and cooperation with users, we found that social incentives also motivate users to work together, understand each other's lifestyle better, and implicitly influence one another. The presence of another one increases users' awareness of both individual performance and social accountability, which is likely to help users engage in their health goals in the long run. Meanwhile, social interaction design should consider involving suitable members in a group, such as with similar exercise goals to guarantee comparability, and within close-knit groups to increase the sense of responsibility. In essence, this chapter goes beyond examining information presentation on the systems to promoting intervention effectiveness through imposing positive group interaction and dynamics.

5 Evaluation Framework

In this chapter, we present an evaluation framework that aims at modeling various interaction strategies, subjective norm, social capital and the influence of these qualities on users' attitudes and behavioral intentions of persuasive technologies for behavior change. We also conduct a small-scale study to validate the model and report the results of statistical tests. The results address the importance of social interaction in social persuasive technologies. Part of this chapter appears in [32].

5.1 Introduction

In the HCI community, there is an increasing interest for developing effective user-centered technological tools for personalized behavior change [42, 94, 83, 128]. However, adopting technology for behavior change is a complex process involving sensitive issues such as the fear of being monitored, the reluctance of receiving recommendations from a device, and privacy concerns [122]. Evaluating a technology's contribution to behavior change and users' acceptance after usage requires large studies and significant resources [122]. Klasnja et al. identified the difficulty of evaluating technologies for health behavior change in HCI field. They argued that focusing on specific strategies (such as self-monitoring) and gaining an in-depth understanding of users' experience of such technologies can be considered feasible evaluation strategies. Due to the high cost in evaluating technologies for behavior change in terms of required resources and time, the literature lacks an evaluation framework that sheds light on the design factors –especially social interfaces and interactions – that influence users' attitudes and acceptance for such technologies.

We aim to investigate how social interaction mechanisms, i.e., **social strategies**, can influence users' attitudes and adoption of technologies for health behavior change. More concretely, we are interested in the following research questions:

- Can **social strategies** be designed to provide positive attitudes and increase users' intention to use technologies for behavior change?

- Can **social factors** (such as subjective norm and perceived social capital) influence users' attitudes and adoption of such technologies?
- Can system **playfulness** influence users' attitudes and usage intention of such technologies?

To answer these questions, we hypothesized a model and evaluated it using factor analysis and correlational tests through an in-lab user survey. This research seeks to understand how social interactions and social influence associate with user acceptance for persuasive technologies for personal healthcare. It also aims to provide practical suggestions to design more efficient pervasive and preventative healthcare systems.

5.2 Related Work

Researchers have investigated and established acceptance models for healthcare systems based on Technology Acceptance Model (TAM) [48]. TAM is a widely accepted intention model for predicting IT usage [19]. It shows that users' *perceived usefulness* (PU) and *perceived ease of use* (PEOU) of a system positively influence their attitudes towards the system, which predict their intention to actual usage. TAM has been extensively used as generic framework to evaluate users' acceptance for technologies. Current studies support TAM as a robust method to model acceptance of IT in health care [70]. However, TAM model has been criticized for ignoring *social* aspects of technology. Alternatively, Unified Theory of Acceptance and Use of Technology (UTAUT) [116] extends the original constructs of PU to *perceived expectancy* (PE), and PEOU to *effort expectancy* (EE) and adds *social influence* (SI) as the third construct that leads to users' intention to use.

A number of researchers have investigated acceptance model for healthcare technologies. Wu et al. [125] examined and evaluated healthcare professionals' usage intention of mobile healthcare systems (MHSs). They have developed a model based on TAM and validated it by survey results collected from 123 healthcare professionals, including physicians, nurses, and medical technicians who were currently using various MHSs. The model indicates that compatibility of MHS devices with users' current technical background and user-perceived ability to use the system significantly predicts PU and PEOU. The Almere [68] model evaluates acceptance of assistive social robots for healthcare of the elderly. It emphasizes on social aspects of a system in user acceptance, such as perceived social presence and perceived sociability. Steele et al. [17] investigated elderly persons' attitudes and acceptance of using wireless sensor networks to assist healthcare. They identified crucial elements that influence acceptance by conducting a focus-group study and showing participants a sample device. Sensor Acceptance Model (SEM) [56] consists of two questionnaires, assessing patients' attitudes before and after using sensor technologies respectively. The pre-study questionnaire (Q1) evaluates patients' physical health, mental health and their expectation for the sensors. The post-study questionnaire (Q2) assesses patients' attitudes towards different dimensions of sensors, such as hygienic aspects, skin reactions, etc. In this model, physical health, mental health and expectations in Q1 influence patients' perception on various dimensions of sensors in Q2 and hence predict sensor acceptance index in Q2.

The above work evaluates technologies designed for *specific user groups*, such as patients, elderly people and healthcare professionals. The technologies considered are designed for their respective needs. For example, elderly people have strong demand for staying independent with the help of technology [68]. Healthcare professionals possess deeper domain knowledge but require some degree of technology support [125]. Patients have more urgent and stronger motivations for the use of technologies to cure or alleviate health conditions that they actually suffer from than average users. Thus far, prior work has not covered technology acceptance issues for *average users* whose routine daily living habits may lead to health concerns. Nor have the above studies investigated the role of *social factors* in user attitudes and adoption of healthcare systems. Our current study aims at understanding and explaining 1) how average users -– non-patients and non-professionals -– accept technologies for preventive care and 2) how social factors influence the acceptance.

5.3 Model Development and Research Hypotheses

Our goal is to understand normal users' acceptance for new technologies that could help them change their behaviors and cultivate healthy habits. To model the social factors, we conceptualized a framework based on Choi and Chung's work which applies Technology Acceptance Model and studies the impact of subjective norm and social capital on the acceptance of social networking services [40]. Their model contains five constructs: *intention to use*, *perceived usefulness*, *perceived ease of use*, *perceived social capital* and *subjective norm*. In the domain of pervasive healthcare, incentives and strategies in systems that motivate users' behavior change play a crucial role in their performance and their attitudes to such systems [80]. Additionally, many of this applications aim to enhance users' enjoyment in using such systems. Therefore, we adapt Choi and Chung's model and integrate another two constructs: *social strategies* and *perceived playfulness*. In this section, we depict research model and hypotheses in Figure 5.1 and explain the constructs and hypotheses.

5.3.1 Behavioral Intention

Behavioral intention refers to an individual's motivation or willingness to exert effort to perform the target behavior [70]. Since behavior-change is a long and complex process [80], one of the fundamental goals for behavior-change systems is to maximize users' retention in using the system and their sustainability of behavior change. Accordingly, this dimension consists of the following criteria: users' intention to use the system again and the intention to introduce this system to her/his friends.

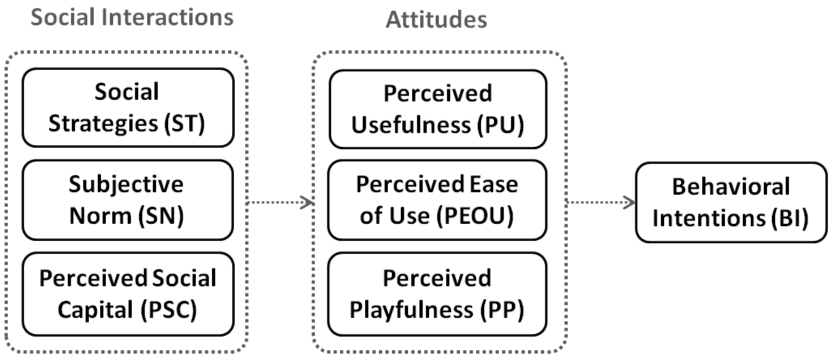


Figure 5.1: Research Model Hypothesis.

5.3.2 Perceived Usefulness

Perceived usefulness is one of the most fundamental factors in the Technology Acceptance Model. Davis defined perceived usefulness as "the degree to which a person believes that using a particular system would enhance his or her job performance." [48] In the context of working, job performance is also evaluated by criteria such as accomplishing tasks more quickly, more efficiently, effectively and easily. For persuasive technologies, users will perceive a system useful if it helps them improve their physical activities or increase their awareness of physical activities.

5.3.3 Perceived Ease of Use

As another fundamental factor in Technology Acceptance Model, Davis defined ***perceived ease of use*** as "the degree to which a person believes that using a particular system would be free of effort." [48] Perceived ease of use is associated with ease and freedom from difficulty or great effort. This criterion also implies the system is clear, understandable, flexible and easily helping people to be skillful in using the system.

5.3.4 Perceived Playfulness

Users' attitudes towards technology are not always rational, and emotion plays an important role [130]. Webster and Martocchio [121] defined micro-computer ***playfulness*** as an individual's tendency to interact spontaneously, inventively, and imaginatively with computers. Some researchers have interchangeably used the terms such as perceived enjoyability [77]. Ahn et al. has studied the impact of playfulness on user acceptance of online retailing [19]. They extended TAM model by integrating playfulness as an "added" variable [70]. Similarly, Padilla-Meléndez et al. [100] identified perceived playfulness as a facilitation factor of perceived usefulness and perceived ease of use in online learning. Their validation results of structural model showed that playfulness significantly affects user attitude and intention to use online retailing services.

5.3.5 Social Strategies

In the domain of pervasive health, researchers have been investigating various social strategies that motivate users to stay physically active. Such strategies include sharing information in social networks, competition, cooperation, communication, and gamification.

Sharing

Many persuasive technologies allow users to share their health data in their social networks. Sharing with families or friends is reported to be effective in motivating users to cultivate healthy habits [60]. This is reflected in the studies of both Consolvo et al. [41] and Toscos et al. [113] that aim to help users enhance physical activities. Ahtinen et al. [20] found that connecting with family members and loved ones can help motivate users to fulfill their fitness goals. Currently, many mobile apps allow users to share their fitness statistics using social media. Typical examples include Nike+, RunKeeper, miCoach, etc. Similar strategies have been applied to help users to manage body-weight [88], form a healthy habit of water-intake [39], and maintain a good sleeping habit [78]. However, Munson and Consolvo [96] pointed out that users may hesitate to share in the large social networks due to the concerns about whom to share and over-sharing. Rather, they would prefer to share with close-knit groups, such as families and close friends.

Competition

Competition is one of the most widely adopted social interaction strategies. Some technologies design competition in a small team, such as Fitster [21], Kukini [29], Fish'n'Steps [85] and Life Coaching Application [59]. Some applications motivate users by enabling them to compete in a large community, such as Nike+ and RunKeeper. The intention of not being the last and aiming for the top does provide some degree of motivation to users, as reported in Chapter 2 and 4.

Cooperation

The cooperation strategy binds users' performance with that of their team members. In the case of persuasive technologies for physical activities, cooperation not only promotes users' exercises but also brings group enjoyment and enhances interpersonal relationship among friends. For example, users of Fish'n'Steps are motivated to enhance their performance for the sake of group responsibility [85]; Chick Clique [112] makes users aware of which members are below average and stimulates them to encourage the teammates who are left behind. We have presented the literature and our comparative studies on the competition and cooperation mechanisms in Chapter 4.

Communication

Communication components in persuasive technologies allow users to interact actively, persuade and influence others. As mentioned in Chapter 2, social communication bring numerous benefits, such as providing a channel to give support to group members [41], raising the awareness of each other's effort [95], increasing a user's responsibility to exercise [22] and enhancing players' intimacy, which builds a solid basis for long-term play experience [29].

Gamification

Gamification component is frequently used in designing persuasive technologies to make sports activities fun. Researchers also identified gamification as a promising factor to motivate users in physical activities [28]. Many applications apply methods such as badges, leaderboards and metaphorical visualizations to encourage individuals to work out more. UbiFit Garden [42] visualizes users' daily steps by the growing status of plants. The more activities a user takes, the healthier his/her plant looks. It also displays butterflies to encourage more varieties of activities. Fish'n'Steps [85] uses the metaphor of fish tank to visualize users' step count. Fitster [21] translates users' accumulated exercise distances to the routes on geographical maps and visualizes exercises as traveling around cities. The social game iFitQuest [87] uses metaphors such as escaping the wolf and collecting coins to train users' speed and endurance in physical exercise.

5.3.6 Subjective Norms

Subjective norms refer to a person's perception that most people who are important to him think he should or should not perform the behavior in question [116] or how important others in his or her social environment wish or expect him or her to behave in a certain way" [40]. In the field of behavior change, subjective norm means the degree to which a person perceives the demands of important or referent others on him or her in the behavior. In Unified Theory of Acceptance and Use of Technology (UTAUT), Venkatesh et al. [116] considered the subjective norm as one aspect of social influence that significantly affects users' behavioral intention to use the system. Choi and Chung [40] have identified the subjective norm as a significant factor that influence users' intention to use social networking sites. In their validated model, subjective norm outweighs perceived usefulness and perceived ease of use among the three determining factors of intention to use social networking sites. This work is relevant with modeling social interfaces because of people's innate propensity of interpersonal bonding with others.

5.3.7 Perceived Social Capital

Perceived social capital refers to users' perception that "the resources are created through interactions in social relationships and provide benefits to participants in the network." [40]

Social capital offers benefits for members in a social network who could obtain resources from other members and leverage connections from multiple social contexts. Choi and Chung [40] further identified two types of social capital. "Bridging social capital refers to the information, or new perspectives gained without personal or emotional experience, usually among loose connections; "bonding social capital" refers to emotional exchanged among close relationships like family or friends. The classification covers the two types of relationships that commonly exist in social network sites. In their validated model, perceived capital significantly influences perceived usefulness and perceived ease of use. We believe the two types of social capital also apply in persuasive technologies for behavior change.

5.4 Model Validation – A Pilot Study

5.4.1 Participants and Procedure

We deployed a pilot study to test the model. We invited 23 users on campus who voluntarily participated in the study, including 12 males and 11 females. Their ages range between 19 and 40. All of them are students except one teacher. They came from 8 different countries, such as Switzerland, France, Ireland, etc. All of them have used Fitbit and HealthyTogether before.

We invited the participants to the laboratory to ensure the quality of their response and to check the accuracy of the wording of the statements. They were instructed to respond to a survey for Fitbit or HealthyTogether based on their personal experience. The survey consists of 24 statements (see Table 5.1). For each statement, a 7-point Likert scale from "strongly disagree" (1) to "strongly agree" (7) was used to characterize users' responses. In the end, we received 46 responses from the participants for Fitbit application and HealthyTogether.

5.4.2 Validity and Reliability of the Constructs

We first compute the internal consistency and reliability of the constructs using Cronbach's alpha and item-total correlations. This validation process aims to check the internal consistency of a given construct, i.e., how closely are items related in a group. Normally, a Cronbach's alpha value higher than 0.70 is considered acceptable [98], but sometimes lower thresholds are accepted. The item-total correlation above the recommended threshold (0.50) is considered acceptable [103]. Table 5.1 shows the values of the constructs. The only item that does not reach the threshold value of item-total correlation is ST5 (the game elements in the system are well designed). One explanation is that game elements can be designed both in individual and social setting. Therefore, it is less correlated with other items that are inherently *social* strategies.

Table 5.1: Test results of internal reliability. Constructs with single item are included for completeness.

Construct	Cronbach	Item-Total Corr.
1. Strategies Klasnja et al. (2011)	.660	
ST1 It is easy to see my friends' steps and stairs using the system.		.475
ST2 It is easy to communicate with my friends using the system.		.528
ST3 It is easy to compete with friends using the system.		.424
ST4 It is easy to cooperate with friends using the system.		.441
ST5 The game elements in the system are well designed.		.082
2. Subjective Norm Choi & Chung (2013)	.794	
SN1 My friends think that more steps and more stairs in the system are important to me.		.395
SN2 It really would not matter to my friends if I decided to give up using the system. (reverse)		.732
SN3 My friends expect me to continuously use the system.		.669
SN4 No one would really be surprised if I just stopped using the system. (reverse)		.644
3. Perceived Social Capital Choi & Chung (2013)	.724	
PSC1 The system make it easier to develop social relationship with my friends.		.725
PSC2 I find the system useful in my social relationship.		.559
PSC3 I feel my friends and I are motivated to have more steps and stairs when using the system.		.296
PSC4 I am concerned about my friends' activities.		.504
4. Perceived Playfulness Tony Ahn et al. (2007)		
PP1 Using HealthyTogether gives enjoyment to me for my task.		
5. Perceived Usefulness Choi & Chung (2013), Davis (1989), Tony Ahn et al. (2007)	.851	
PU1 The system increases my motivation to do physical exercise.		.733
PU2 Using the system improves my overall activity volume.		.740
PU3 I feel more connected with my buddy and friend circles after using the system.		.702
PU4 I feel more familiar with my buddy and friend circles after using the system.		.503
PU5 I find the system useful in my daily life.		.651
6. Perceived Ease of Use Choi & Chung (2013), Davis (1989)	.813	
PEOU1 My interaction with the system is clear and understandable.		.630
PEOU2 Interacting with the system does not require a lot of my mental effort.		.677
PEOU3 I find the system easy to use.		.685
7. Behavioral Intentions Choi & Chung (2013), Pu et al. (2011), Tony Ahn et al. (2007)	.727	
BI1 I will continue to use the system in the future.		.614
BI2 I will tell my friends about the system.		.614

Table 5.2: Pearson correlations between behavioral intentions and user attitudes towards a system (*$p < .05$, **$p < .01$).

	BI1 (Continue to use)	**BI2**(Tell friends)
PP1	.701** (0)	.662** (0)
PU1	.595** (0)	.498** (0)
PU3	.338* (.022)	.382** (.009)
PU5	.718** (0)	.492** (.001)
PEOU3	.431** (.003)	.409** (.005)

5.4.3 Correlational Tests

We then conducted correlational analysis based on our hypothesized model. Since it is a small-scale (N = 46) test of the model, we used correlational analysis instead of inferential analysis [98]. The correlation coefficients and the corresponding statistical significance (p-value) do not indicate casual relationship, but to what extent two variables are associated.

Behavioral Intentions

We first investigate how *behavioral intentions* – including the intention to use (BI1) and tell friends about the system (BI2) – are associated with users' attitudes towards the system, namely *perceived playfulness, perceived usefulness* and *perceived ease of use*. Note that we mainly examined three aspects of perceived usefulness: 1) usefulness in improving performance (PU1&PU2), 2) usefulness in enhancing social relationship (PU3&PU4) and 3) overall system usefulness (PU5). When designing the questionnaire, we intentionally repeat some items to cross-validate survey answers [105]. However, due to the small sample size, we only choose one item of each aspect for correlational tests. We select PU1, PU3 and PU5 as items for the construct of 'perceived usefulness'. Similarly, we select PEOU3 as representative item of perceived ease of use in correlational test. We listed the correlational analysis in Table 5.2. Results show that users' usage intention is significantly related with perceived playfulness (PP1) ($r = .701, p = 0$), perceived usefulness in improving performance (PU1) ($r = .595, p = 0$), perceived usefulness in enhancing social relationship ($r = .338, p = .022$) (PU3), overall perceived usefulness (PU5) ($r = .718, p = 0$), and perceived ease of use (PEOU3) ($r = .431, p = .003$).

Users' intention to tell friends about the system is strongly associated with perceived playfulness (PP1) ($r = .662, p = 0$), perceived usefulness in performance (PU1) ($r = .498, p = 0$), perceived usefulness in social relationship (PU3) ($r = .382, p = .009$), overall perceived usefulness (PU5) ($r = .492, p = .001$), and perceived ease of use (PEOU3) ($r = .405, p = .005$).

The above results confirm with the models that were built on top of technology acceptance model. Particularly, perceived playfulness, compared with perceived ease of use, is more closely correlated with willingness to continue to use. Noticeably, perceived playfulness has

Table 5.3: Pearson correlations among perceived playfulness, perceived usefulness and perceived ease of use (*$p < .05$, **$p < .01$).

	PP	PU1	PU3	PU5	PEOU3
PU1	.686** (0)				
PU3	.521** (0)	.524** (0)			
PU5	.584** (0)	.669** (0)	.406** (.005)		
PEOU3	.370* (.011)	0.245 (.10)	0.133 (.379)	.512** (0)	1

the highest association with willingness to tell friends about the system, compared with the other factors. The above results confirm the previous findings that system playfulness can enhance users' retention of the systems [19].

User Attitudes

We then examine the relationship among the three constructs of users' attitudes towards a system: *perceived playfulness, perceived usefulness* and *perceive ease of use.* As explained earlier, we select PP1, PU1, PU3, PU5 and PEOU3 to present these three constructs in correlational tests. The results (see Table 5.3) show that perceived playfulness (PP1), overall perceived usefulness (PU5) and perceived ease of use (PEOU3) significantly associate with each other. The above results again confirm with the TAM model.

However, we have not found any significant correlation between perceived usefulness in enhancing physical activities (PU1) and perceived ease of use (PEOU3) ($r = .245, p = .10$). Nor have we found the significance between perceived usefulness in enhancing social relationship (PU3) and perceived ease of use (PEOU3) ($r = .133, p = .379$). One explanation is that users' performance in their physical activities and social relationship are not directly influenced by how easy it is to use the systems. The results hint that designing strategies and incentives are important in user-perceived usefulness and social relationship for behavioral change.

We then investigate how *perceived ease of use* is associated with *social strategies, social norm* and *perceived social capital.* Among all the items in the constructs, we only identified two items significantly correlated with perceived ease of use. First, it is positively correlated with gamification (ST5) – the game elements in the system are well designed ($r = .300, p = .043$). Additionally, the level of perceived ease of use is associated with the reverse scale of "people would care if I stop using the system" (reverse scale of SN4, r = .310, p = .036). Thus, we consider gamification and subjective norm as important elements in enhancing user-perceived ease of use.

We further examine how *perceived usefulness* and *perceived playfulness* associate with *social strategies, social norm* and *perceived social capital.* We list all *significant* correlations in Table 5.4. Among all the social strategies, perceived system usefulness and playfulness are only significantly associated with gamification (ST5). By contrast, perceived usefulness

Table 5.4: Pearson correlations between perceived usefulness, perceived playfulness and social strategies, social capital and subjective norm (*$p < .05$, **$p < .01$).

	PU5	**PP1**
ST5	.340* (.021)	.256 (.086)
SN1	.229 (.126)	.29 (.051)
SN2 (rev)	.291* (.049)	.293* (.048)
SN3	.412** (.004)	.16 (.287)
SN4 (rev)	.336* (.023)	.288 (.053)
PSC1	.360* (.014)	.499** (0)
PSC2	.267 (.073)	.383** (.009)
PSC4	.347* (.018)	.300* (.043)

Table 5.5: Pearson correlations between strategies and subjective norm (*$p < .05$, **$p < .01$).

	SN1	**SN2** (rev)	**SN3**	**SN4** (rev)
ST1	.115 (.448)	-.002 (.989)	.029 (.85)	-.01 (.945)
ST2	**.429** (.003)**	-.091 (.547)	.214 (.153)	.088 (.562)
ST3	**.286 (.054)**	-.058 (.701)	.185 (.217)	.101(.506)
ST4	**.530** (0)**	**.334* (.023)**	**.277 (.062)**	0.231 (.122)
ST5	.086 (.569)	-.094 (.533)	.023 (.879)	-0.112 (.459)

is significantly correlated with both user perceived social norm and social capital. The same result is found for perceived playfulness. This hints that when users perceive they can gain benefits out of social interactions, they are more likely to consider the system useful and playful.

Social Strategies

Finally, we investigate whether the social norm is correlated with the various social strategies. Table 5.5 lists the correlations among these factors. SN1 refers to the social norm for one's daily physical activities while SN2, SN3, SN4 focus on one's system usage. SN1 is strongly associated with social communication component (r = .429, p = .003), competition (r = .286, p = .054), and cooperation (r = .530, p = 0). This implies the motivating effects of communication component, competition and cooperation of users' physical activities. Interestingly, neither communication nor competition has a significant correlation with the social norm in using the system (SN2, SN3, and SN4). By contrast, this correlation is close to significant for cooperation. The above results hint that designing social strategy that support users to cooperate with others help them to sustain in using the system.

5.5 Summary

In this chapter, we propose an evaluation framework for persuasive technologies. In particular, this model aims to explore the impact of *social strategies* and *social influence* in users' attitudes and their usage intention. Moreover, this model emphasizes the role of perceived playfulness in influencing users' behavioral intentions, in addition to perceived usefulness and perceived ease of use that are covered in the classical Technology Acceptance Model (TAM). We further deploy a small-scale pilot study to assess the framework. We collect 46 data entries via an in-lab survey for pervasive fitness applications. The results of factor analysis and correlational tests provide a preliminary validation for our theoretical framework. In particular, our results show that perceived system playfulness is strongly correlated with users' intention to continue to use the system and to tell their friends about the system. Additionally, designing appropriate social strategies, user-perceived social capital and subjective norm are positively correlated with users' attitudes of the system (e.g., perceived usefulness and perceived playfulness).

We are aware of the limitation of the small sample size in the validation of the study. It restricts us from drawing conclusions on causal relationships and conducting regression analysis such as structural equation modeling (SEM). In the future, we will conduct a large-scale study that covers more varieties of persuasive technologies to further validate the model.

6 Design Guidelines

The previous chapters cover social interface and interaction design guidelines for groupware (Chapter 2) and group recommenders (Chapter 3). They are evaluated with users, the findings of which provide lessons on *whether* the design works, *why* it works, and *what* are the success factors. Each design guideline is motivated by a *requirement* in enhancing user experience in the social context. It is then supported with some *techniques* and *practices* to meet the requirement. In this chapter, we summarize the design guidelines that aim to enhance the following aspects of social experience of group recommender systems: mutual awareness, intimacy, immediacy, group dynamics, motivation, and playfulness.

6.1 Mutual Awareness

The ***mutual awareness*** feature – well-informed knowledge among group members – has been extensively investigated in social software. As discussed in previous chapters, most social interfaces mainly help users to be aware of each other on information for *very specific tasks*. This makes them insufficient to support full awareness in relation to the offline social interactions. Some researchers have offered some solutions to address the "partial awareness" issues. For example, Meeting Mediator [79] detects and visualizes group dynamics in group meeting. Group dynamic, which seems irrelevant with group tasks, is demonstrated to help users improve their task performance and cooperation cohesion [79]. Other works provide awareness of prosody, facial movement and gestures in group working space [102]. The above work creates awareness features that *appear to be* irrelevant to the collaboration tasks. However, studies show that these features improved users' performance and experience in collaboration. In fact, the essence of these works is not to create new information to visualize, but to compensate the invisible and subtle cues that are easy to capture in offline social interactions. We refer to such cues as ***side-channel information***.

As mentioned in Chapter 4, mutual awareness features in group recommender systems mainly cover membership awareness, preference awareness and decision awareness [30, 31]. They provide basic information needed to generate group recommendations. We then go further to integrate emotion awareness features by designing emotional interfaces. Study results show that being aware of group members' emotions towards recommended items, although not directly needed for generating recommendations, help enhance group members' intimacy, perception of the interaction immediacy, and influence their attitudes towards recommendations [33]. Therefore, we propose the following design guideline:

Design Principle 6.1.1 (Awareness)**.** *When designing group recommenders, consider compensating awareness features for side-channel information that are pervading in offline social interaction; such information includes emotion, prosody, facial expressions, etc.*

Pentland has criticized the "*social ignorance*" of most groupware [102]. Failing to transmit side-channel information is one cause of this issue. Another factor worth noting is the insufficiency of the *transmission channel*. In offline social interactions, people perceive the environment using their five organs – one not only seeing, but also hearing, smelling, tasting, and touching in the social context. While graphical user interface is the most conventional modality to transmit information, much context is lost due to lacking involvement of other sense organs and thus causes the problem of "*information incompleteness.*" Chapter 2 covers systems that provide not only visual format, but also audio and haptic ways. Therefore, we propose the following design guideline:

Design Principle 6.1.2 (Awareness)**.** *Consider providing multiple modalities of information channels to enhance the level of awareness among group members during collaboration; multiple modalities cover not only graphical interfaces, but also speech, haptic, olfactory, etc.*

Pervasive sensors have played a crucial role in informing side-channel cues and enabling multi-modality interfaces. However, it does not mean overwhelming users with excessive information and making awareness feature cumbersome. Design Guideline 2.4.9 and 2.4.7, we suggest:

Design Principle 6.1.3 (Awareness)**.** *When designing awareness features, consider using lightweight input method and balancing between information richness and user effort.*

6.2 Intimacy

Intimacy – or relatedness, closeness [66] – is another indispensable expectation in social interaction. In Chapter 2, we summarize techniques that aim to enhance intimacy among loved ones [66], couples [24], close-knit groups [45], family members [124], and colleagues [85, 39]. The first technique to enhance intimacy is providing awareness of *seemingly trivial information.* In face-to-face social interaction, closeness is commonly achieved through being aware of myriad types of cues about each other, ranging from preference in music and food to lifestyles and daily routines. Some seemingly insignificant information about the group members can help them know each other better. For example, CoupleVibe helps couples who live apart to be aware of one another's location [24], which triggers the users to frequently think about and imagine each other's life. BuddyClock facilitates users to care about friends' sleeping schedule in a small group and accommodate decisions in when to contact each other [78].

In Chapter 3, we present how emotional interfaces can enhance users' intimacy with each other in group recommender systems [33]. In Chapter 4, we discover that designing appropriate social incentives in behavior change can also effectively increase group members' intimacy [36, 38]. In HealthyTogether, users reported that being aware of buddy's steps help them perceive their friends' overall lifestyle and daily activities, which leads to a higher level of familiarity [38]. Having a better knowledge of friends' music preferences also promotes their intimacy [33]. Thus, we recommend:

Design Principle 6.2.1 (Intimacy)**.** *Do not neglect mutual awareness of seemingly insignificant information in enhancing intimacy.*

The second technique is emotion awareness. Emotions are natural responses of one's attitudes to an internal or external states, as opposed to reasoning. Audiences in a pop concert spontaneously dance to the music with joy or energy; people laugh out loud after watching fun episodes of video clips; tourists naturally give rise to the feeling of awe in front of wondrous landscape or scenery. Emotions are considered important channels in expressing one's attitudes and in social communication. The study results of Cui et al. [45] reveal that recording users' facial expressions as emotional feedback to friends' photos can increase the intimacy among group members and lead to greater group enjoyment. Visualizing friends' emotional responses while listening to music also help users know about each other's music tastes and increase the group satisfaction in group music recommender system [33]. Thus, we propose the following design guideline:

Design Principle 6.2.2 (Intimacy)**.** *Consider using emotional feedback to enhance the perception of intimacy among non-collocated group members.*

Shared activity is another commonly used technique to help establish a feeling of relatedness. Life experience shows that shared activities, such as playing games, traveling, exercising together, can help friends build positive memory, and establish a closer relationship. In this aspect, the fact that group members are bonded for an activity is more important than the activity itself. For example, jogging together in different locations help users transform running from a personal experience to an interpersonal one [95]. CoDine allows family members who are apart to eat "together" mediated by pervasive display and sensors [124]. Achieving health goals together also demonstrates to be effective in enhancing intimacy. For example, HealthyTogether users have reported organizing working-out together to gain more steps and floors [38, 37]. So we propose:

Design Principle 6.2.3 (Intimacy)**.** *Consider establishing shared activities to transform personal experience to interpersonal interaction.*

6.3 Immediacy

Immediacy refers to "the quality that makes something seem important or interesting because it is or seems to be happening *now*", according to the Merriam-Webster Dictionary [2]. In online communication, the goal of increasing users' perception of immediacy is to reduce the perceived distance between people in distributed locations [46]. The task is further complicated by the situation for asynchronous communication, when multiple parties are not interacting at the same time.

One pattern to produce immediacy for online communication is to present information with temporal cues. Jogging Over a Distance [95] is designed for distributed running partners to jog together. It continuously senses users' heart rate and translates it to spatial audio in relation to users' location (i.e., in front, aside or behind). The temporal feedback about exercising partner's running status shortens their perceived distance. This is a typical example of synchronous social experience. Temporal cues are also widely adopted in asynchronous communication. Social Camera [45] allows users to provide feedback to their friends' photos by recording their first-response facial expressions, which are deemed as emotional responses. Instead of providing textual comments, using static emoticons or "like" buttons, instant videos provide temporal cues that make users feel as if the feedback is provided immediately.

In Chapter 3, we introduce empatheticons in group recommender, which visualize users' emotional responses to music by animating their profile pictures to the timeline of the song. Study results show that users perceive themselves as listening to music together with other group members, even though the study was deployed in an asynchronous setting. Thus, we propose the following design guideline:

Design Principle 6.3.1 (Immediacy)**.** *Consider presenting temporal information to reduce perceived physical distance in online social interaction; such temporal cues could enhance the immediacy of both synchronous and asynchronous online social interaction.*

Physicality is an indispensable factor in mediating connectedness during offline social interaction. Typical examples include hugging, kissing, holding, etc. CoupleVibe employs vibrations to convey location meanings between couples [24]. The design of CoDine also integrates tangible artifacts such as printing edible text messages from co-diners while dining to enhance co-experience [124]. Hassenzahl et al. [66] have surveyed techniques that aim to create or enhance the experience of relatedness and consider physicality to be important yet challenging to implement. Wang et al. [120] investigated how to influence the perception of connectedness using mediated remote social touch to convey emotional experiences. Their study demonstrated that the communication accompanied by touch modality results in a significant increase in the sense of connectedness and decrease in the perceived physical distance. Thus, we suggest:

Design Principle 6.3.2 (Immediacy)**.** *Consider physicality to increase the feeling of connectedness and shorten perceived physical distance in distributed group interaction.*

6.4 Group Dynamics

Group dynamics are defined as “the influential interpersonal processes that take place in groups." [58] A group is not simply the sum of its members. Instead, the members interact with and influence each other, and change their attitudes and behaviors dynamically. BuddyClock [78] visualizes sleeping status among a group of friends. By being aware of each other’s patterns of sleeping schedule, users already appeared to form more aligned schedules for sleeping over a period of two weeks. Similar phenomenon occurred in the study results of Playful Bottle [39], which show that the amount of daily water-intake tends to be closer for participants in a group in four weeks.

Our empatheticon study also shows that users’ emotional feedback when listening to music is significantly correlated with that of other group members. The number of users’ emotional annotations increases with that of their group members. Since this invisible influence naturally exists among social groups, it is desirable to design for positive influence. Thus, we propose:

Design Principle 6.4.1 (Group Dynamics)**.** *Consider enforcing positive group influence to enhance group consensus.*

Even though group dynamics inevitably exist within a group, people may not always be aware of them. It is possible to intervene this process to enhance group performance by visualizing them. Meeting Mediator [79] detects group members’ speech, body movement, proximity and face-to-face interaction and analyzes and visualizes their activeness and dominance in a group. Results show that being aware of the group dynamics facilitates collaboration and improves the efficiency for group tasks such as problem-solving and brainstorming. We thus recommend:

Design Principle 6.4.2 (Group Dynamics)**.** *Consider intervening group dynamics by visualizing them to enhance group performance in collaboration.*

6.5 Motivation

One challenging goal of designing social interfaces is to motivate members to contribute. In group recommender systems, it is desirable to avoid “the rich getting richer and the poor getting poorer" problem – a phenomenon that occurs when only a few active members contribute while others remain lurkers. In the case of group recommender systems, one research question is how to motivate users to indicate their preferences in the group actively. Persuasive technologies for behavior change also aim to motivate all group members to participate vigorously. The results of Playful Bottle show that users who are monitoring their daily water-intake in a group setting drank significant more water than those in an individual setting [39]. Similar results were found in the study of HealthyTogether that users exercising in a group were more motivated than when they walked alone [37]. In the long run, users in the social environment are more likely to sustain in the behavior change process. This is also shown in the study results of VERA [25]. Placing users in a social setting serves as a motivation for behavior change. Thus, we suggest:

Design Principle 6.5.1 (Motivation)**.** *Consider making users aware of each other's progress and performance of behavior change in a group or social community to enhance user performance and sustainability in the long run.*

Researchers have integrated various social interaction schemes in persuasive technologies for behavior change. Users may compete with others [85], cooperate to reach exercise goals [41], be accountable for others' daily steps [38], persuade buddies to take water [39], or encourage each other by sending virtual gifts [39]. The effectiveness of these social interactions may vary in different settings [37]. Designing social interaction strategies are likely to increase users' perception of social capital. Therefore, we derive the following principle:

Design Principle 6.5.2 (Motivation)**.** *Consider designing social incentives to afford different types of group interactions, such as cooperation, competition, social accountability, persuasion, and gift-giving.*

Communication is an essential component in social interfaces. A well-designed communication mechanism can greatly enhance users' experience and engagement in a group. In this thesis, we cover various communication components, such as transforming partners' heart-rate to spatial sound during jogging [95], conveying location information using vibrations [24], delivering persuasion and care via gift-giving [39], expressing emotional feedback to a music with empatheticons [33], etc. Chiu et al. also found that persuasions received from friends for behavior change were more efficient than those obtained from the machine [39]. However, messaging components should be designed in a user-friendly way to scaffold users to send messages. Typical techniques include providing templates, emoticons, etc. Based on the discussions, we provide the following design guideline:

Design Principle 6.5.3 (Motivation)**.** *Consider affording user-friendly communication components to motivate users to interact with and persuade each other; such communication may be implemented in various modalities and scaffold users' interaction.*

Meanwhile, several studies have showed that users' motivation varies with who the group members are. For example, while some users like competition in fitness applications, they reported competing with a much stronger or weaker buddy could de-motivate them [37]. For another example, users were not motivated by cooperation in a team if they did not know with whom they were working together [85, 25]. So we suggest:

Design Principle 6.5.4 (Motivation)**.** *Consider helping users find suitable buddies to motivate their behavior change.*

6.6 Playfulness

Playfulness, as defined by Webster and Martocchio, refers to an individual's tendency to interact spontaneously, inventively, and imaginatively with computers [121]. Chapter 5 has extensively studied playfulness and shows that playfulness is positively associated with users' perception of the usefulness of social fitness applications. The most commonly used technique for enhancing playfulness is gamification – the use of game elements in non-game context [51]. Games are inherently fun; applying them in a social context naturally introduces another layer of enjoyment when interacting with others, which is one major success factor of social games such as FarmVille [4] and DrawSomething [3]. Applying social gamification elements in applications for behavior change also improves the performance. Study results of Playful Bottle indicate that users had healthier water-drinking habits in game settings than individual settings, and social games outperform individual games [39]. HealthyTogether studies also show that when users are playing a social game, they have more fun and more activities in the intervention process [37]. Thus we suggest:

Design Principle 6.6.1 (Playfulness)**.** *Consider designing social gamification strategies in applications to create playfulness and increase intervention effectiveness in behavior change.*

However, it is worth noting that social game elements, such as badges, points, competition, are extrinsic motivations for behavior change. The ultimate goal is to motivate users intrinsically, which is very difficult. Studies in Chapter 4 show that after the intervention phase, some users' physical activities have increased compared with before intervention. Therefore, we consider social gamification strategies as an expedient and effective method for the early stage of intervention.

6.7 Summary

This chapter summarizes social interface design guidelines derived from the thesis. These guidelines cover six requirements: mutual awareness, intimacy, immediacy, group dynamics, motivation, and playfulness. They can help researchers and practitioners enhance the social user experience of the system. The guidelines do not cover *all* areas of social requirements. After all, offline social interaction is a complex process. For example, the flip side of awareness, intimacy and immediacy is privacy. From the users' perspective, the main privacy concern involves sharing information with strangers [25]. Another issue is to avoid over-sharing information beyond the minimum requirement [24]. We do not extensively discuss privacy issues in this thesis since it is another research area. Apropos, the guidelines *do* address important – yet easily neglected – aspects of user experience in the social context, rather than providing a solution that satisfying *all* requirements.

7 Conclusions

Group recommender systems help a group of users find items of interest. Recommendation may be generated from the system proactively making suggestions or from the members persuading and influencing each other. However, the current solutions tend to emphasize on the algorithm aspect of the systems and lack the attention on users' social experience. This thesis aims to investigate how to design for interaction and experience among users mediated by computers, which we refer to as social interfaces. In this thesis, we study social interface and interaction design for group recommenders systems. We first survey the user issues in group recommender systems. In order to understand the best interface and interaction design practice, we present ten typical social applications that highlight the design and implementation of interfaces and derive a preliminary set of guidelines. Based on these guidelines, we have developed two experimental platforms – *GroupFun*, a music group recommender system and *HealthyTogether*, a group fitness application that supports social persuasion. Using GroupFun, we have studied emotion awareness tools as social interfaces for group recommenders. We have designed three emotion representation methods and empirically demonstrated the effectiveness of emotion awareness in promoting social interaction and interpersonal relationship in group recommender systems. With HealthyTogether, we have extensively investigated social interaction schemes. We have created various social interaction mechanisms in motivating users to exercise, such as social accountability, competition, cooperation and social capital. Our three user studies that span over a period of up to ten weeks with 80 users provide empirical evidence that designing and adjusting social incentives plays a significant role in users' performance. We also propose an evaluation framework for persuasive technologies for behavior change. The framework aims to establish the relationship among social interaction design of an application, users' perceptions of the social context, their attitudes towards the system, and their behavioral intentions. Finally, we single out six criteria for designing social interfaces and interactions to enhance interpersonal interaction and users' group experience with the systems. We derive a total of fifteen design guidelines for these criteria.

To the best of our knowledge, this thesis is the first work that extensively examines the design and effectiveness of social interfaces and interactions in influencing user experience in group recommenders. In the following sections, we detail the contribution in each aspect of the work. We then point out the limitations of the thesis and future research directions.

7.1 Contributions to HCI

7.1.1 Design

Social Interface Design for Group Recommenders

We have surveyed user issues in group recommender systems and identified the lack of emphasis on emotion awareness. To fill this gap, we first developed a music group recommender system called GroupFun as an experimental platform for emotion awareness tools. We designed and produced the visualization techniques of color wheels (for ACTI and CoFeel) and dynamic animations (for empatheticons). We then conducted a user study to evaluate the roles and effectiveness of emotion awareness feature in group recommender systems by implementing empatheticons in GroupFun. Our results show that empatheticons can help users familiarize with other group members' preferences, enhance their perception of immediacy in non-collocated communication, promote their interpersonal relationship, and have the potential to enforce positive group dynamics. Our work offers multiple contributions. First, we present the design of novel emotion awareness tools in online group environment; second, we empirically demonstrate the roles of emotion awareness in creating positive group dynamics in group recommender systems; third, we uncover how emotions – *seemingly* irrelevant information – enhance the social glue in an online collaborative environment.

Social Interaction Design for Group Recommenders

We study social interaction design for group recommender systems using persuasive technologies. We developed a mobile application called HealthyTogether that allows dyads to walk together and win badges through different rewarding mechanisms. We used HealthyTogether as an experimental platform for three controlled user studies that compare users' physical activities when they exercise individually and with buddies. The first study, with 24 participants, compares the effectiveness of applying social accountability, competition and a hybrid model as rewarding mechanisms in motivating users' physical exercise. The second, with 36 participants, explores three social incentives: competition, cooperation and a hybrid of them. The third, a longitudinal study with 20 users, examines the role of teamwork in goal-setting, badge system and leaderboard. In all three studies, social settings (i.e., when users walk with a buddy) outperformed individual setting (i.e., when users walk alone). Furthermore, we report the following findings. First, social accountability and cooperation help users enhance the interpersonal relationship with buddies; second, competition with other people may have the risk of de-motivating users, and thus selecting a comparable buddy is critical; third, users' motivation in exercising is strongly associated with the amount of communication between the buddies; finally, walk in a team to fulfill exercise goals motivate users in the long run. Our user studies, involving 80 users over a period of up to ten weeks, are the first work that extensively explores the motivational effects of group interaction in persuasive technologies for physical exercise. The results manifest how employing appropriate group interaction schemes can enhance users' performance in behavior change. They also lead to strong implications of designing social interaction schemes for persuasive technology.

7.1.2 Evaluation

We also propose an evaluation framework that aims to identify and assess the factors that affect users' attitudes and behavioral intention for social pervasive applications. The framework aims to examine how various social incentives, subjective norm and perceived social capital influence users' perception of the systems' playfulness, usefulness and ease of use, and ultimately impact their behavioral intention for the system. The model is assessed with a small-scale study (N=46) through correlational analysis. Our preliminary results show that users' behavioral intention with the system is strongly associated with the perceived system playfulness, which is positively correlated with their perception of the system's usefulness and ease of use. Results also suggest that the perceived playfulness is related with the design of social strategies. The social strategy design has a significant association with user-perceived social capital. This framework theoretically pictures the importance of designing social interaction schemes in users' willingness to use the system and recommend it to their social networks. The results of the model validation lead to design implications for social interfaces and interactions to enhance users' acceptance of persuasive technologies.

7.1.3 Guidelines

We derive fifteen guidelines for designing social interfaces and interaction for group recommenders, based on the survey of social interface design for groupware, findings of our user studies, and the evaluation framework. These guidelines are summarized to fulfill the following six requirements that we have identified to enhance users' social experience: mutual awareness, intimacy, immediacy, group dynamics, motivation, and playfulness. To improve the level of mutual awareness, it is essential to provide awareness for side-channel information, present it in multiple modalities, but require minimum or lightweight input from users. To enhance users' intimacy with each other, we suggest providing mutual awareness of seemingly insignificant information, emotions, or shared activities. For immediacy property, we propose integrating temporal cues and physicality in interface design. Additionally, we suggest visualizing group influence or enforcing it in a positive way to facilitate group consensus and improve intervention effect. As for motivation, we recommend designing appropriate social interaction schemes, communication components, and matching proper buddies to reinforce behavior change. Finally, we emphasize on designing social gamification strategies to enhance the playfulness and intervention effects for behavior change. These guidelines can help designers and developers to design more social-friendly applications.

7.2 Limitations

This thesis also suffers from a few limitations. For the design, we only investigate social interfaces through the lens of two domains: group recommender systems for music and behavior change. Even though they are representative areas, we are aware that the lessons we learned from the design should be adapted accordingly when applying in other domains. For the evaluation, we propose a generic framework for persuasive technology for behavior change. The framework is assessed through correlational test. Further inferential analysis through structural equation modeling (SEM) would validate the causal relationship among the constructs, once we can obtain data from more users using more varieties of systems [105]. For the design guidelines, admittedly, we can not and do not claim to cover all design aspects for social interfaces, e.g., privacy, but so far we have addressed the most comprehensive areas that are easily neglected for social interactions.

7.3 Future Work

In this section, we point out a few research directions that will extend our studies and findings.

7.3.1 Social Incentive Design for Patient Care

This thesis demonstrates the effectiveness of designing social incentives to motivate users' physical exercises. The subjects of the studies are mainly recruited from campus, most of whom have a sedentary lifestyle. Encouraged by the positive results, we envision that such applications could benefit patients with chronic diseases such as obesity and diabetes, which are leading causes of cardiovascular diseases and cancer. The World Health Organization has recommended physical activities as one of the crucial factors to prevent obesity and diabetes [9], but it is considered challenging to introduce patients to form exercise habits into sedentary lifestyles [111]. Presumptuously, social applications such as HealthyTogether should have similar intervention effectiveness for patients with chronic diseases. However, the effects may be even more promising since patients are more in need of technologies to stimulate their physical activities, or they may expect more motivation mechanisms such as engaging them in the long term due to the chronicity of their disease. In the future, we are interested in validating our findings with the patients and identifying potentially more specific requirements for patient care.

7.3.2 Social Interface Design for Emotional Well-being

Staying physically active is one healthy goal. According to The World Health Organization, health also includes other dimensions such as mental and social well-being [15]. According to the Stress In America 2013 report published by the American Psychological Association, the urge for stress management for both teenagers and adults is on the rise [14]. Social science also suggests the importance of improving mental wellbeing by using the social networks. Besides fitness applications, we are also interested in applying the guidelines derived from the thesis to design, develop and evaluate applications for emotional well-being. We are interested in understanding whether providing social support, such as making group members aware of each other's emotions can enhance their emotional well-being. Chapter 3 presents three types of emotion-awareness tools. We are motivated to apply these tools in social applications for emotional well-being.

7.3.3 Integrated Evaluation Framework

The third research direction is to continue with validating the evaluation framework. Currently, we evaluate the proposed framework in a small-scale user survey involving two social fitness applications. In line of this direction, we plan to conduct a large-scale user study to assess the validity and reliability of the model using structured equation modeling (SEM). While extending the scale of the study, we also aim to check the generalizability of the model [105] by recruiting users who have experienced other social fitness applications on their own, such as Jawbone, Nike+, Runtastic, RunKeeper, etc. After revising and evaluating the current model, we intend to investigate how to generalize the model for social interface design in group recommender systems and other application domains.

7.4 Take-home Messages

In short, this thesis presents the design, evaluation and guidelines of social interfaces and interactions for group recommender systems. Particularly, it extensively discusses two domains: group recommender that support online collaboration using the case of music group recommender system, and those involve group influence and persuasion using the case of pervasive fitness applications. The effectiveness and social roles of the interfaces are evaluated with an in-depth lab study with 24 users and three field studies with 80 users for up to ten weeks. The thesis then presents an evaluation framework that aims to show how social interaction strategies, social capital and subjective norm are associated with users' attitudes and acceptance for social fitness applications. Based on the literature and our study results, we derive fifteen social design guidelines and address six main criteria to enhance users' social experience which tend to be neglected: mutual awareness of side-channel information, intimacy, immediacy, group dynamics, motivation and playfulness.

Social applications, which are in the larger context of group recommenders, are pervasive in our daily life. This thesis sheds light on the user issues of such applications – how to reduce the perceived presence of the computers and increase the perceived connections with people on the other side of the computers. As Nicholas Negroponte said, "*Computing is not about computers any more. It is about living.*" Thus, the main target is to offer group-centered interfaces instead of system functionalities, to address the influence of users' perceptions rather than information or reasoning. Essentially, these guidelines aim to provoke and draw designers' attention on "leaky holes of a pipeline" when reconstructing social interactions online.

Bibliography

[1] CityVille. http://zynga.com/game/cityville. [Online; accessed 8-Aug-2014].

[2] Definition of Immediacy in Merriam Webster Dictionary. http://www.merriam-webster.com/dictionary/immediacy. [Online; accessed 8-Aug-2014].

[3] DrawSomething. http://zynga.com/game/drawsomething. [Online; accessed 8-Aug-2014].

[4] FarmVille. http://zynga.com/game/farmville. [Online; accessed 8-Aug-2014].

[5] Fitbit. https://www.fitbit.com/. [Online; accessed 8-Aug-2014].

[6] GoalSponsor. https://www.goalsponsors.com. [Online; accessed 8-Aug-2014].

[7] miCoach. http://micoach.adidas.com/. [Online; accessed 8-Aug-2014].

[8] Nike+. https://secure-nikeplus.nike.com/plus/. [Online; accessed 8-Aug-2014].

[9] Obesity. http://www.who.int/topics/obesity/en/. [Online; accessed 21-July-2014].

[10] Pareto Principle. http://en.wikipedia.org/wiki/Pareto_principle. [Online; accessed 8-Aug-2014].

[11] RunKeeper. http://runkeeper.com/. [Online; accessed 8-Aug-2014].

[12] Runtastic. https://www.runtastic.com/. [Online; accessed 8-Aug-2014].

[13] Stickk.com. http://www.stickk.com/. [Online; accessed 8-Aug-2014].

[14] Stress in America. http://www.apa.org/news/press/releases/stress/2013/highlights.aspx. [Online; accessed 21-July-2014].

[15] WHO definition of Health. http://www.who.int/about/definition/en/print.html. [Online; accessed 21-July-2014].

[16] Zynga Inc. http://zynga.com/. [Online; accessed 8-Aug-2014].

[17] Elderly persons' perception and acceptance of using wireless sensor networks to assist healthcare. *International Journal of Medical Informatics 78*, 12 (2009), 788 – 801.

[18] ADAMS, P. *Grouped: How small groups of friends are the key to influence on the social web.* New Riders, 2011.

[19] AHN, T., RYU, S., AND HAN, I. The impact of web quality and playfulness on user acceptance of online retailing. *Information & Management 44*, 3 (2007), 263 – 275.

[20] AHTINEN, A., ISOMURSU, M., MUKHTAR, M., MÄNTYJÄRVI, J., HÄKKILÄ, J., AND BLOM, J. Designing social features for mobile and ubiquitous wellness applications. In *Proceedings of the 8th International Conference on Mobile and Ubiquitous Multimedia* (2009), MUM '09, ACM, pp. 12:1–12:10.

[21] ALI-HASAN, N., GAVALES, D., PETERSON, A., AND RAW, M. Fitster: Social fitness information visualizer. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (2006), CHI EA '06, ACM, pp. 1795–1800.

[22] ANDERSON, I., MAITLAND, J., SHERWOOD, S., BARKHUUS, L., CHALMERS, M., HALL, M., BROWN, B., AND MULLER, H. Shakra: Tracking and sharing daily activity levels with unaugmented mobile phones. *Mobile Networks and Applications 12*, 2-3 (2007), 185–199.

[23] ARDISSONO, L., GOY, A., PETRONE, G., SEGNAN, M., AND TORASSO, P. Intrigue: Personalized recommendation of tourist attractions for desktop and hand held devices. *Applied Artificial Intelligence 17*, 8-9 (2003), 687–714.

[24] BALES, E., LI, K. A., AND GRIWSOLD, W. Couplevibe: Mobile implicit communication to improve awareness for (long-distance) couples. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (2011), CSCW '11, ACM, pp. 65–74.

[25] BAUMER, E., KHOVANSKAYA, V., ADAMS, P., POLLAK, J., VOIDA, S., AND GAY, G. Designing for engaging experiences in mobile social-health support systems. *Pervasive Computing 12*, 3 (July 2013), 32–39.

[26] BERKOVSKY, S., AND FREYNE, J. Group-based recipe recommendations: Analysis of data aggregation strategies. In *Proceedings of the Fourth ACM Conference on Recommender Systems* (2010), RecSys '10, ACM, pp. 111–118.

[27] BRADLEY, M. M., AND LANG., P. J. Affective reactions to acoustic stimuli. *Psychophysiology 37*, 02 (2000), 204–215.

[28] BRAUNER, P., CALERO VALDEZ, A., SCHROEDER, U., AND ZIEFLE, M. Increase physical fitness and create health awareness through exergames and gamification. In *Human Factors in Computing and Informatics*, vol. 7946 of *Lecture Notes in Computer Science*. 2013.

[29] CAMPBELL, T., NGO, B., AND FOGARTY, J. Game design principles in everyday fitness applications. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (2008), CSCW '08, ACM, pp. 249–252.

[30] CHEN, Y. Interface and interaction design for group and social recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (2011), RecSys '11, ACM, pp. 363–366.

[31] CHEN, Y. User issues in social group recommender systems. Tech. rep., EPFL, 2011.

[32] CHEN, Y., LE, D., YUMAK, Z., AND PU, P. EHR: A technology acceptance model for lifestyle changes. (journal article in preparation).

[33] CHEN, Y., MA, X., CEREZO, A., AND PU, P. Empatheticons: Designing emotion awareness tools for group recommenders. In *Proceedings of the XV International Conference on Human Computer Interaction* (2014), Interaccion'14, ACM.

[34] CHEN, Y., AND PU, P. Do you feel how we feel? an affective interface in social group recommender systems. In *Proceedings of the 3rd Workshop on Recommender Systems and the Social Web* (2011), RSWEB '11.

[35] CHEN, Y., AND PU, P. Cofeel: Using emotions for social interaction in group recommender systems. In *Proceedings of the First International Workshop on Interfaces for Recommender Systems* (2012), InterfaceRS '12), pp. 48–55.

[36] CHEN, Y., AND PU, P. Designing emotion awareness interface for group recommender systems. In *Proceedings of the 2014 International Working Conference on Advanced Visual Interfaces* (2014), AVI '14, ACM, pp. 347–348.

[37] CHEN, Y., AND PU, P. Healthytogether: Exploring social incentives for mobile fitness applications. In *Proceedings of the Second International Symposium of Chinese CHI* (2014), Chinese CHI '14, ACM, pp. 25–34.

[38] CHEN, Y., ZHANG, J., AND PU, P. Exploring social accountability for pervasive fitness apps. In *Proceedings of the Eighth International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies* (2014), Ubicomm'14.

[39] CHIU, M.-C., CHANG, S.-P., CHANG, Y.-C., CHU, H.-H., CHEN, C. C.-H., HSIAO, F.-H., AND KO, J.-C. Playful bottle: A mobile social persuasion system to motivate healthy water intake. In *Proceedings of the 11th International Conference on Ubiquitous Computing* (2009), Ubicomp '09, ACM, pp. 185–194.

[40] CHOI, G., AND CHUNG, H. Applying the technology acceptance model to social networking sites (sns): Impact of subjective norm and social capital on the acceptance of sns. *International Journal of Human-Computer Interaction 29*, 10 (2013), 619–628.

[41] CONSOLVO, S., EVERITT, K., SMITH, I., AND LANDAY, J. A. Design requirements for technologies that encourage physical activity. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006), CHI '06, ACM, pp. 457–466.

[42] CONSOLVO, S., MCDONALD, D. W., TOSCOS, T., CHEN, M. Y., FROEHLICH, J., HARRISON, B., KLASNJA, P., LAMARCA, A., LEGRAND, L., LIBBY, R., SMITH, I., AND LANDAY, J. A. Activity sensing in the wild: A field trial of ubifit garden. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2008), CHI '08, ACM, pp. 1797–1806.

[43] CROSSEN, A., BUDZIK, J., AND HAMMOND, K. J. Flytrap: Intelligent group music recommendation. In *Proceedings of the 7th International Conference on Intelligent User Interfaces* (2002), IUI '02, ACM, pp. 184–185.

[44] CRUMLISH, C., AND MALONE, E. *Designing social interfaces: Principles, patterns, and practices for improving the user experience.* O'Reilly Media, Inc., 2009.

[45] CUI, Y., KANGAS, J., HOLM, J., AND GRASSEL, G. Front-camera video recordings as emotion responses to mobile photos shared within close-knit groups. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), CHI '13, ACM, pp. 981–990.

[46] D. BAKER, J. An investigation of relationships among instructor immediacy and affective and cognitive learning in the online classroom. *The Internet and Higher Education 7*, 1 (2004), 1 – 13.

[47] DALLERY, J., CASSIDY, R. N., AND RAIFF, B. R. Single-case experimental designs to evaluate novel technology-based health interventions. *Journal of medical Internet research 15*, 2 (2013).

[48] DAVIS, F. D. Perceived usefulness, perceived ease of use, and user acceptance of information technology. *MIS quarterly* (1989), 319–340.

[49] DERKS, D., BOS, A. E., AND VON GRUMBKOW, J. Emoticons and social interaction on the internet: the importance of social context. *Computers in Human Behavior 23*, 1 (2007), 842 – 849.

[50] DESMET, P. Measuring emotion: Development and application of an instrument to measure emotional responses to products. In *Funology*, M. Blythe, K. Overbeeke, A. Monk, and P. Wright, Eds., vol. 3 of *Human-Computer Interaction Series*. Springer Netherlands, 2005, pp. 111–123.

[51] DETERDING, S., SICART, M., NACKE, L., O'HARA, K., AND DIXON, D. Gamification. using game-design elements in non-gaming contexts. In *CHI '11 Extended Abstracts on Human Factors in Computing Systems* (2011), CHI EA '11, ACM, pp. 2425–2428.

[52] DEY, A. K., AND DE GUZMAN, E. From awareness to connectedness: The design and deployment of presence displays. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2006), CHI '06, ACM, pp. 899–908.

[53] EKMAN, P. Basic emotions. In *Handbook of Cognition and Emotion*. John Wiley and Sons, Ltd.

[54] ELLIS, C. A., GIBBS, S. J., AND REIN, G. Groupware: Some issues and experiences. *Commun. ACM 34*, 1 (Jan. 1991), 39–58.

[55] FAN, C., FORLIZZI, J., AND DEY, A. K. A spark of activity: Exploring informative art as visualization for physical activity. In *Proceedings of the 2012 ACM Conference on Ubiquitous Computing* (2012), UbiComp '12, ACM, pp. 81–84.

[56] FENSLI, R., P.-P. E. G. T. . H. O. Technology acceptance; assistive technology; elderly users; social robots; embodied agents. *Methods Inf Med 47* (2008), 89–95.

[57] FORLIZZI, J., AND BATTARBEE, K. Understanding experience in interactive systems. In *Proceedings of the 5th Conference on Designing Interactive Systems: Processes, Practices, Methods, and Techniques* (2004), DIS '04, ACM, pp. 261–268.

[58] FORSYTH, D. *Group dynamics*. Thomson Higher Education, 2009.

[59] GASSER, R., BRODBECK, D., DEGEN, M., LUTHIGER, J., WYSS, R., AND REICHLIN, S. Persuasiveness of a mobile lifestyle coaching application using social facilitation. In *Persuasive Technology*, vol. 3962 of *Lecture Notes in Computer Science*. 2006.

[60] GRIMES, A., TAN, D., AND MORRIS, D. Toward technologies that support family reflections on health. In *Proceedings of the ACM 2009 International Conference on Supporting Group Work* (2009), GROUP '09, ACM, pp. 311–320.

[61] GROSS, J. J., AND LEVENSON, R. W. Emotion elicitation using films. *Cognition and Emotion 9*, 1 (1995), 87–108.

[62] GRUDIN, J. Groupware and social dynamics: Eight challenges for developers. *Commun. ACM 37*, 1 (Jan. 1994), 92–105.

[63] GUTWIN, C., SCHNEIDER, O., XIAO, R., AND BREWSTER, S. Chalk sounds: The effects of dynamic synthesized audio on workspace awareness in distributed groupware. In *Proceedings of the ACM 2011 Conference on Computer Supported Cooperative Work* (2011), CSCW '11, ACM, pp. 85–94.

[64] HALKO, S., AND KIENTZ, J. Personality and persuasive technology: An exploratory study on health-promoting mobile applications. In *Persuasive Technology*, vol. 6137 of *Lecture Notes in Computer Science*. 2010.

[65] HARRISON, C., HSIEH, G., WILLIS, K. D., FORLIZZI, J., AND HUDSON, S. E. Kineticons: Using iconographic motion in graphical user interface design. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), CHI '11, ACM, pp. 1999–2008.

[66] HASSENZAHL, M., HEIDECKER, S., ECKOLDT, K., DIEFENBACH, S., AND HILLMANN, U. All you need is love: Current strategies of mediating intimate relationships through technology. *ACM Trans. Comput.-Hum. Interact. 19*, 4 (Dec. 2012), 30:1–30:19.

[67] HATT, J. A. H. The colorist. *Journal of the Royal Society of Arts 57*, 2973 (1909), 1025 – 1025.

[68] HEERINK, M., KRÖSE, B., EVERS, V., AND WIELINGA, B. Assessing acceptance of assistive social agent technology by older adults – the almere model. *International Journal of Social Robotics 2*, 4 (2010), 361–375.

[69] HERR, S., RÖSCH, A., BECKMANN, C., AND GROSS, T. Informing the design of group recommender systems. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (2012), CHI EA '12, ACM, pp. 2507–2512.

[70] HOLDEN, R. J., AND KARSH, B.-T. The technology acceptance model: Its past and its future in health care. *Journal of Biomedical Informatics 43*, 1 (2010), 159 – 172.

[71] HUISMAN, G., VAN HOUT, M., VAN DIJK, E., VAN DER GEEST, T., AND HEYLEN, D. Lemtool: Measuring emotions in visual interfaces. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), CHI '13, ACM, pp. 351–360.

[72] IJSSELSTEIJN, WIJNAND, J. V. B., AND VAN LANEN, F. Staying in touch: Social presence and connectedness through synchronous and asynchronous communication media. *Human-Computer Interaction: Theory and Practice (Part II) 2* (2003), 924–928.

[73] JAIMOVICH, J., COGHLAN, N., AND KNAPP, R. Emotion in motion: A study of music and affective response. In *From Sounds to Music and Emotions*, vol. 7900 of *Lecture Notes in Computer Science*. 2013.

[74] JAMESON, A. More than the sum of its members: Challenges for group recommender systems. In *Proceedings of the Working Conference on Advanced Visual Interfaces* (2004), AVI '04, ACM, pp. 48–54.

[75] JAMNER, M. S., SPRUIJT-METZ, D., BASSIN, S., AND COOPER, D. M. A controlled evaluation of a school-based intervention to promote physical activity among sedentary adolescent females: project {FAB}. *Journal of Adolescent Health 34*, 4 (2004), 279 – 289.

[76] JOHANSEN, R. *GroupWare: Computer Support for Business Teams.* The Free Press, 1988.

[77] JONES, N., AND PU, P. User technology adoption issues in recommender systems. *NAEC, ATSMA* (2007), 379–39.

[78] KIM, S., KIENTZ, J. A., PATEL, S. N., AND ABOWD, G. D. Are you sleeping?: Sharing portrayed sleeping status within a social network. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (2008), CSCW '08, ACM, pp. 619–628.

[79] KIM, T., CHANG, A., HOLLAND, L., AND PENTLAND, A. S. Meeting mediator: Enhancing group collaborationusing sociometric feedback. In *Proceedings of the 2008 ACM Conference on Computer Supported Cooperative Work* (2008), CSCW '08, ACM, pp. 457–466.

[80] KLASNJA, P., CONSOLVO, S., AND PRATT, W. How to evaluate technologies for health behavior change in hci research. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2011), CHI '11, ACM, pp. 3063–3072.

[81] KOVECSES, Z. *Metaphor and emotion: Language, culture, and body in human feeling.* Cambridge University Press, 2003.

[82] KUDENKO, D., BAUER, M., AND DENGLER, D. Group decision making through mediated discussions. In *User Modeling 2003*, vol. 2702 of *Lecture Notes in Computer Science.* 2003.

[83] LI, I., DEY, A., AND FORLIZZI, J. A stage-based model of personal informatics systems. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2010), CHI '10, ACM, pp. 557–566.

[84] LI, I., DEY, A. K., AND FORLIZZI, J. Understanding my data, myself: Supporting self-reflection with ubicomp technologies. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (2011), UbiComp '11, ACM, pp. 405–414.

[85] LIN, J., MAMYKINA, L., LINDTNER, S., DELAJOUX, G., AND STRUB, H. Fish'n'steps: Encouraging physical activity with an interactive computer game. In *UbiComp 2006: Ubiquitous Computing* (2006), vol. 4206 of *Lecture Notes in Computer Science.*

[86] LO, S.-K. The nonverbal communication functions of emoticons in computer-mediated communication. *CyberPsychology and Behavior 11*, 5 (2008), 595 — 597.

[87] MACVEAN, A., AND ROBERTSON, J. Understanding exergame users' physical activity, motivation and behavior over time. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2013), CHI '13, ACM, pp. 1251–1260.

[88] MAITLAND, J., AND CHALMERS, M. Finding a balance: Social support v. privacy during weight-management. In *CHI '08 Extended Abstracts on Human Factors in Computing Systems* (2008), CHI EA '08, ACM, pp. 3015–3020.

[89] MARK, G., IQBAL, S., CZERWINSKI, M., AND JOHNS, P. Capturing the mood: Facebook and face-to-face encounters in the workplace. In *Proceedings of the 17th ACM Conference on Computer Supported Cooperative Work and Social Computing* (2014), CSCW '14, ACM, pp. 1082–1094.

[90] MASTHOFF, J., AND GATT, A. In pursuit of satisfaction and the prevention of embarrassment: affective state in group recommender systems. *User Modeling and User-Adapted Interaction 16*, 3-4 (2006), 281–319.

[91] McCarthy, J. F. Pocket restaurantfinder: A situated recommender system for groups. In *Workshop on Mobile Ad-Hoc Communication at the 2002 ACM Conference on Human Factors in Computer Systems* (2002).

[92] McCarthy, J. F., and Anagnost, T. D. Musicfx: An arbiter of group preferences for computer supported collaborative workouts. In *Proceedings of the 1998 ACM Conference on Computer Supported Cooperative Work* (1998), CSCW '98, ACM, pp. 363–372.

[93] McCarthy, K., Salamó, M., Coyle, L., McGinty, L., Smyth, B., and Nixon, P. Group recommender systems: A critiquing based approach. In *Proceedings of the 11th International Conference on Intelligent User Interfaces* (2006), IUI '06, ACM, pp. 267–269.

[94] McDuff, D., Karlson, A., Kapoor, A., Roseway, A., and Czerwinski, M. Affectaura: An intelligent system for emotional memory. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (New York, NY, USA, 2012), CHI '12, ACM, pp. 849–858.

[95] Mueller, F., Vetere, F., Gibbs, M., Edge, D., Agamanolis, S., Sheridan, J., and Heer, J. Balancing exertion experiences. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), CHI '12, ACM, pp. 1853–1862.

[96] Munson, S., and Consolvo, S. Exploring goal-setting, rewards, self-monitoring, and sharing to motivate physical activity. In *Pervasive Computing Technologies for Healthcare (PervasiveHealth), 2012 6th International Conference on* (May 2012), pp. 25–32.

[97] Neviarouskaya, A., Prendinger, H., and Ishizuka, M. Emoheart: Conveying emotions in second life based on affect sensing from text. *Adv. in Hum.-Comp. Int. 2010* (Jan. 2010).

[98] Nunally, J. C., and Bernstein, I. H. *Psychometric theory.* McGraw Hill.

[99] O'Connor, M., Cosley, D., Konstan, J., and Riedl, J. Polylens: A recommender system for groups of users. In *ECSCW 2001*. 2001.

[100] Padilla-Meléndez, A., Aguila-Obra, A. R. d., and Garrido-Moreno, A. Perceived playfulness, gender differences and technology acceptance model in a blended learning scenario. *Computers and Education 63*, 0 (2013), 306 – 317.

[101] Peng, W., and Hsieh, G. The influence of competition, cooperation, and player relationship in a motor performance centered computer game. *Computers in Human Behavior 28*, 6 (2012), 2100 – 2106.

[102] Pentland, A. Socially aware, computation and communication. *Computer 38*, 3 (March 2005), 33–40.

[103] Peterson, R. A. A meta-analysis of cronbach's coefficient alpha. *Journal of Consumer Research 21*, 2, pp. 381–391.

[104] POPESCU, G., AND PU, P. What's the best music you have?: Designing music recommendation for group enjoyment in groupfun. In *CHI '12 Extended Abstracts on Human Factors in Computing Systems* (2012), CHI EA '12, ACM, pp. 1673–1678.

[105] PU, P., CHEN, L., AND HU, R. A user-centric evaluation framework for recommender systems. In *Proceedings of the Fifth ACM Conference on Recommender Systems* (2011), RecSys '11, ACM, pp. 157–164.

[106] PU, P., AND POPESCU, G. Eye-tracking group influence-experiment design and results. EPFL technical report.

[107] RIVERA, K., COOKE, N. J., AND BAUHS, J. A. The effects of emotional icons on remote communication. In *Conference Companion on Human Factors in Computing Systems* (1996), CHI '96, ACM, pp. 99–100.

[108] RUSSELL, J. A. A circumplex model of affect. *Journal of Personality and Social Psychology 39*, 6 (1980), 1161– 1178.

[109] SCHERER, K. R. What are emotions? and how can they be measured? *Social science information 44*, 4 (2005), 695 – 729.

[110] SCHLICHTER, J., KOCH, M., AND BÜRGER, M. Workspace awareness for distributed teams. In *Coordination Technology for Collaborative Applications*, vol. 1364 of *Lecture Notes in Computer Science*. 1998.

[111] SMITH, B., FROST, J., ALBAYRAK, M., AND SUDHAKAR, R. Integrating glucometers and digital photography as experience capture tools to enhance patient understanding and communication of diabetes self-management practices. *Personal and Ubiquitous Computing 11*, 4 (2007), 273–286.

[112] TOSCOS, T., FABER, A., AN, S., AND GANDHI, M. P. Chick clique: Persuasive technology to motivate teenage girls to exercise. In *CHI '06 Extended Abstracts on Human Factors in Computing Systems* (2006), CHI EA '06, ACM, pp. 1873–1878.

[113] TOSCOS, T., FABER, A., CONNELLY, K., AND UPOMA, A. Encouraging physical activity in teens can technology help reduce barriers to physical activity in adolescent girls? In *the Second International Conference on Pervasive Computing Technologies for Healthcare* (Jan 2008), PervasiveHealth '08, pp. 218–221.

[114] TRAN, M. H., YANG, Y., AND RAIKUNDALIA, G. K. Supporting awareness in instant messaging: An empirical study and mechanism design. In *Proceedings of the 17th Australia Conference on Computer-Human Interaction: Citizens Online: Considerations for Today and the Future* (2005), OZCHI '05, pp. 1–10.

[115] TU, C.-H. The measurement of social presence in an online learning environment. *International Journal on E-Learning 1*, 2 (2002), 34–45.

[116] VENKATESH, V., MORRIS, M. G., DAVIS, G. B., AND DAVIS, F. D. User acceptance of information technology: Toward a unified view. *MIS Quarterly 27*, 3 (2003), pp. 425–478.

[117] VETERE, F., SMITH, J., AND GIBBS, M. Phatic interactions: Being aware and feeling connected. In *Awareness Systems*, Human-Computer Interaction Series. 2009.

[118] WALTHER, J. B., AND D'ADDARIO, K. P. The impacts of emoticons on message interpretation in computer-mediated communication. *Social science computer review 19*, 3 (2001), 324 – 347.

[119] WANG, H., PRENDINGER, H., AND IGARASHI, T. Communicating emotions in online chat using physiological sensors and animated text. In *CHI '04 Extended Abstracts on Human Factors in Computing Systems* (2004), CHI EA '04, ACM, pp. 1171–1174.

[120] WANG, R., QUEK, F., TATAR, D., TEH, K. S., AND CHEOK, A. Keep in touch: Channel, expectation and experience. In *Proceedings of the SIGCHI Conference on Human Factors in Computing Systems* (2012), CHI '12, ACM, pp. 139–148.

[121] WEBSTER, J., AND MARTOCCHIO, J. J. Microcomputer playfulness: development of a measure with workplace implications. *MIS quarterly* (1992), 201–226.

[122] WEEKS, J. W., HEIMBERG, R. G., FRESCO, D. M., HART, T. A., TURK, C. L., SCHNEIER, F. R., AND LIEBOWITZ, M. R. Empirical validation and psychometric evaluation of the brief fear of negative evaluation scale in patients with social anxiety disorder. *Psychological assessment 17*, 2 (2005), 179.

[123] WEI, J., CHEOK, A. D., AND NAKATSU, R. Let's have dinner together: Evaluate the mediated co-dining experience. In *Proceedings of the 14th ACM International Conference on Multimodal Interaction* (2012), ICMI '12, ACM, pp. 225–228.

[124] WEI, J., WANG, X., PEIRIS, R. L., CHOI, Y., MARTINEZ, X. R., TACHE, R., KOH, J. T. K. V., HALUPKA, V., AND CHEOK, A. D. Codine: An interactive multi-sensory system for remote dining. In *Proceedings of the 13th International Conference on Ubiquitous Computing* (2011), UbiComp '11, ACM, pp. 21–30.

[125] WU, J.-H., WANG, S.-C., AND LIN, L.-M. Mobile computing acceptance factors in the healthcare industry: A structural equation model. *International Journal of Medical Informatics 76*, 1 (2007), 66 – 77.

[126] YU, Z., ZHOU, X., HAO, Y., AND GU, J. Tv program recommendation for multiple viewers based on user profile merging. *User Modeling and User-Adapted Interaction 16*, 1 (2006), 63–82.

[127] YU, Z., ZHOU, X., AND ZHANG, D. An adaptive in-vehicle multimedia recommender for group users. In *Vehicular Technology Conference* (May 2005), vol. 5, pp. 2800–2804.

[128] YUMAK, Z., AND PU, P. Survey of sensor-based personal wellness management systems. *BioNanoScience 3*, 3 (2013), 254–269.

[129] ZENTNER, M., GRANDJEAN, D., AND SCHERER, K. R. Emotions evoked by the sound of music: Characterization, classification, and measurement. *Emotion 8*, 4 (2008), 494–521.

[130] ZHANG, P., AND LI, N. The importance of affective quality. *Commun. ACM 48*, 9 (Sept. 2005), 105–108.

[131] ZHAO, S. Toward a taxonomy of copresence. *Presence: Teleoperators and Virtual Environments 12*, 5 (2003), 445–455.

Appendices

A Experiment 1

Goal Investigate the roles of effectiveness of empatheticons (Section 3.6)

Duration 1 hour

Participants $N = 18$

Incentives A music-theme USB stick for each participant.

Procedure

1. Before study
   (a) schedule one hour with each participant via email
   (b) obtain their Facebook IDs
   (c) create groups for the participants
   (d) prepare a list of songs for each group
2. During study
   (a) inform the goal, procedure and scenarios of the study to participants
   (b) help them to familiarize with GroupFun and empatheticons with a warm-up song
   (c) listen to 10 songs recommended for the group and annotate emotions using empatheticons
   (d) observe and take notes during the experiment
3. After study
   (a) instruct participants to fill in the post-study survey

Survey Questions

1. Overall, the emotion interface has successfully visualized the emotions in a musical context.
2. The emotion interface is useful in GroupFun.
3. The emotion interface is easy to use.
4. The emotion interface is easy to learn.
5. The emotion interface is novel.
6. The emotion interface is entertaining to use.
7. I immediately felt my friends' emotions.
8. I felt I was listening to music with my friends.
9. I paid close attention to my friends' emotions while listening to music.
10. My emotion annotations were influenced by those of my friends.
11. The songs recommended by GroupFun fit my tastes.
12. The songs recommended by GroupFun fit my friends' tastes.
13. I am satisfied with group experience with GroupFun.
14. I would like to use GroupFun again in the future given the opportunity.

B Experiment 2

Goal Compare competition and accountability for pervasive fitness apps (Section 4.3)

Duration 6 days – 2 weeks

Participants $N = 24$

Incentives A FNAC gift voucher with a value of 40 CHF

Procedure

1. Before study
   (a) schedule 30 minutes with participants via email
   (b) debrief the goal and procedure of the study
   (c) participants sign informed consent
   (d) distribute Fitbit and demonstrate how to set up account
2. Study Phase I
   (a) participants use Fitbit for one week (or 3 days)
   (b) participants complete a daily diary
   (c) at the end of this phase, invite participants to the lab to distribute mobile phones and install HealthyTogether
3. Study Phase II
   (a) participants use HealthyTogether with a buddy for one week (or 3 days)

   (b) participants complete a daily diary

4. After study

   (a) conduct and audio-record a semi-structured post-study interview

Interview Questions

1. What do you think HealthyTogether is?
2. How do you like the rule of the game? Any comments?
3. Can you share your experience with Fitbit?
4. Can you share your experience with HealthyTogether?
5. You used it because of novelty or you will use again?
6. Given the opportunity, would you go back and use it again?
7. What is your motivation in using HealthyTogether?
8. Do you know each other well? Any changes after the study?
9. Any concerns for the game?
10. Any suggestions or comments?

Daily Diary See Figure C.1.

C Experiment 3

Goal Compare competition and cooperation for pervasive fitness apps (Section 4.4)

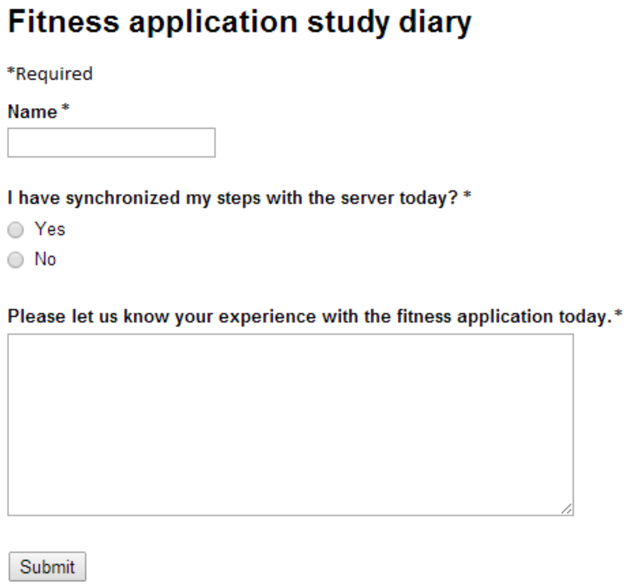
Duration 2 weeks

Participants $N = 36$

Incentives A FNAC gift voucher with a value of 40 CHF

Procedure

1. Before study
   (a) schedule 30 minutes with participants via email
   (b) debrief the goal and procedure of the study
   (c) participants sign informed consent
   (d) distribute Fitbit and demonstrate how to set up account
   (e) let participants to explore Fitbit for two days as *warm-up*
2. Study Phase I
   (a) participants use Fitbit for one week
   (b) participants complete a daily diary
   (c) at the end of this phase, invite participants to the lab to distribute mobile phones and install HealthyTogether
3. Study Phase II



Fitness application study diary

*Required

Name *

I have synchronized my steps with the server today? *

Yes

No

Please let us know your experience with the fitness application today. *

Submit

Figure C.1: Screenshot of daily diary.

   (a) participants use HealthyTogether with a buddy for one week
   (b) participants complete a daily diary

4. After study
   (a) conduct an unstructured post-study interview

Interview Questions Unstructured interview asking about users' experience.

Daily Diary See Figure C.1.

D Experiment 4

Goal Investigate social capital for pervasive fitness apps (Section 4.5)

Duration 10 weeks

Participants $N = 20$

Incentives Fitbit *or* a FNAC gift voucher with a value of 50 CHF, at their own choice

Procedure

1. Before study
   (a) schedule 30 minutes with participants via email
   (b) debrief the goal and procedure of the study and have them sign informed consent
   (c) conduct a semi-structured interview to survey their background
   (d) distribute Fitbit and demonstrate how to set up account
2. Study Phase I
   (a) participants use Fitbit for four weeks
   (b) at the end of this phase, invite participants to the lab and conduct a 30-minute semi-structured interview to survey their 4-week experience with Fitbit
3. Study Phase II
   (a) participants use HealthyTogether with a buddy for four weeks
   (b) at the end of this phase, invite participants to the lab and conduct a 30-minute semi-structured interview to survey their 4-week experience with HealthyTogether

4. Study Phase III
   (a) participants use Fitbit for two weeks

Interview Questions questions covered in the three interviews:

Interview 1

1. Self-introduction
   (a) name, age, nationalities, occupation, residence place, hobbies, etc
2. Social life
   (a) Do you live alone or share with others?
   (b) Do you belong to an association?
   (c) Do you often meet friends?
3. Daily activities
   (a) How you move you from one place to another?
   (b) When do you usually wake up? When do you usually arrive home? When do you usually go to sleep?
4. Physical activity
   (a) Do you practice sports? How often and How long is each session?
   (b) If not often rarely do you, what prevents you to practice regularly?
   (c) Do you exercise in a group or alone? Why?
5. Technological background and attitudes
   (a) Do you have a smartphone? When did you start to use it?
   (b) Any changes before and after you started to use your smartphone?
6. Background about fitness sensors
   (a) Have you ever heard of sensing technology for health management? If yes, under what circumstances?
   (b) Do you know someone who has already used it? If yes, what did he say about his experience?
   (c) Have you ever used such tools? If yes, can you tell us about your experience?
   (d) What do you expect from such sensors or devices?

Interview 2

1. What do you think of Fitbit?
   (a) Do you like it? Do you think it is useful? Is it easy to use?
   (b) Have you met any problems using it?
   (c) Did you ever forget to wear Fitbit? Have you ever lost it?
2. What do you think about the interface of Fitbit?
   (a) Do you use the mobile phone or the computer to check your data?
   (b) How do you think about the dashboard? Is it easy to understand?
   (c) Do you like the way it delivers information?
   (d) Would you like to see other information?
3. Please talk about your experience using Fitbit
   (a) Do you look at the screen of Fitbit frequently?
   (b) How often do you charge?
   (c) What do you think of the messages you have received?
   (d) Where do you usually wear it?
   (e) Did you have any problem synchronizing the data?
   (f) How often do you check your data every day? When do you do it?
4. What were the constraints you had?
   (a) Did you feel discomfort when wearing it ?
   (b) What is the environment when you wear it? People already noticed it?
   (c) Do you think it is accurate in illustrating the effort you spent.
5. Have you noticed change in your behavior ?
   (a) Did you walk more?
   (b) Did you done any physical exercise with Fitbit?
   (c) Have you observed changes in the way of travel?
   (d) How many minutes would you walk every day?
6. Any changes in physical/mental condition ?
   (a) Are you happy using it? Do you plan to continue with it after this study?
   (b) Have you been frustrated by Fitbit?
   (c) Do you think Fitbit has motivated you?

7. What do you think of exercise in group ?
   (a) Have you exercised in pairs or groups ?
   (b) Do you work with a buddy could motivate you to do more exercise? If yes, how would you interact with it through the application?
   (c) Do you think the competition would be a motivation for you? If so, why? If not, why?

Interview 3

1. Experience about HealthyTogether
   (a) What do you think of HealthyTogether in general?
   (b) What do you think of the interface of HealthyTogether?
      i. What do you think about earning badges together with your buddy? Any stories?
      ii. What do you think of the common goal with your buddy? Any stories?
      iii. What do you think of the leaderboard? Any stories to share?
   (c) How often do you use HealthyTogether?
   (d) When do you usually think of using HealthyTogether?
   (e) Any constraints you have with HealthyTogether?
   (f) Have you noticed any changes in your behavior?
2. Now compare the Fitbit app and HealthyTogether app.
   (a) Both have badges, what do you think of them?
   (b) Both have goal setting, what do you think of them?
   (c) Both have leaderboard, what do you think of them?

Yu CHEN

Ph.D. Candidate, Human Computer Interaction

Contact

Room 145, Building BC, EPFL
Station 14, 1015 Lausanne, Switzerland

Tel: +41 21 693 1326, Mobile: +41 789 219 719
http://hci.epfl.ch/members/yu-chen/

Research Interest

I design, develop and evaluate novel interfaces and interaction methods in social media systems. My current research interest focuses on novel emotion-aware interfaces in online social environment. I also investigate interaction methods that motivate behavior change using social influence theories. The goal of my dissertation is to derive a set of social interface design guidelines, which I plan to apply in the broader areas of social media systems.

Education

Ph.D. Candidate, Human Computer Interaction 2010-present
Swiss Federal Institute of Technology at Lausanne (EPFL), Lausanne, Switzerland
Advisor: Dr. Pearl Pu

M.S. Security and Mobile Computing 2009–2010
Norwegian University of Science and Technology, Trondheim, Norway

M.S. Security and Mobile Computing 2008–2010
Aalto University, Helsinki, Finland.

B.E. Information Security 2004–2008
Huazhong University of Science and Technology, Wuhan, China

Research Experience

Nokia Research Center, Helsinki 2009-2010
Research intern. Mentor: Dr. Zheng Yan

Book Chapters

Zheng Yan, Valtteri Niemi, **Yu Chen**, Peng Zhang and Raimo Kantola (2014). Towards Trustworthy Mobile Social Networking. In *Mobile Social Networking* (pp. 195-235). Springer New York

Journals

Zheng Yan, **Yu Chen** and Yue Shen (2014). PerContRep: A Practical Reputation System for Pervasive Content Services. *Journal of Supercomputing* (pp. 1-24). Springer US

Zheng Yan, **Yu Chen** and Yue Shen (2013). A Practical Reputation System for Pervasive Social Chatting. *Journal of Computer and System Sciences*, 79(5) (pp. 556-572)

Conference and Workshops

Yu Chen, Jiyong Zhang and Pearl Pu (in press). Exploring Social Accountability for Pervasive Fitness Apps. In *8th International Conference on Mobile Ubiquitous Computing, Systems, Services and Technologies* (Ubicomm)

Yu Chen, Xiaojuan Ma, Alfredo Cerezo and Pearl Pu (in press). Empatheticons: Designing Emotion Awareness Tools for Group Music Experience. In *15th International Conference on Human Computer Interaction* (Interaccion)

Yu Chen and Pearl Pu (2014). CoFeel: Designing Emotion Awareness Tools for Group Recommender Systems. In *AVI 2014 International Working Conference on Advanced Visual Interfaces* (pp. 347-348). ACM

Yu Chen and Pearl Pu (2014). HealthyTogether: Designing Social Incentives for Mobile Fitness Applications. In *2nd International Symposium of Chinese CHI 2014* (pp. 25-34). ACM

Zheng Yan, **Yu Chen** and Peng Zhang (2012). An Approach of Secure and Fashionable Recognition for Pervasive Face-to-face Social Communications. In *IEEE 8th International Conference on Wireless and Mobile Computing, Networking and Communications* (WiMob), 2012 (pp. 853-860). IEEE

Yu Chen and Pearl Pu (2012). CoFeel: Using Emotions for Social Interaction in Group Recommender Systems. In *1st International Workshop on Interfaces for Recommender Systems* (InterfaceRS 2012) (pp. 48-55)

Yu Chen and Zheng Yan (2012). Gemini: A Handbag for Pervasive Social Communications. In *2012 IEEE 11th International Conference on Trust, Security and Privacy in Computing and Communications* (TrustCom) (pp. 820-825). IEEE

Zerrin Yumak-Kasap, **Yu Chen** and Pearl Pu (2012) EmoSoNet: An Emotion-aware Social Network for Emotional Wellbeing. In *SIGCHI 2012 Workshop on Interaction Design and Emotional Wellbeing*

Yu Chen, Zheng Yan and Valtteri Niemi (2011). Implementation of A Reputation System for Pervasive Social Networking. In *2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications* (TrustCom), (pp. 857-862). IEEE

Yu Chen (2011). Interface and Interaction Design for Group and Social Recommender Systems. In *the 5th ACM conference on Recommender systems* (pp. 363-366). ACM

Yu Chen and Pearl Pu (2011). Do You Feel How We Feel? An Affective Interface in Social Group Recommender Systems. In *the 3rd Workshop on Recommender Systems and the Social Web*

Zheng Yan and **Yu Chen** (2011). AdChatRep: A Reputation System for MANET Chatting. *In the 1st International Symposium on From Digital Footprints to Social and Community Intelligence* (pp. 43-48). ACM

Zheng Yan and **Yu Chen** (2010). AdContRep: A Privacy Enhanced Reputation System for MANET Content Services. In *Ubiquitous Intelligence and Computing* (pp. 414-429). Springer Berlin Heidelberg

Skills

Programming: Java, PHP, C, C++, Python, Mysql
Platforms: Android, Maemo (Nokia Operating Systems), Linux
Statistical packages: SPSS, Matlab, R, AMOS
Research methods: wireframe, survey design and implementation, experimental design, data analysis, SEM, qualitative research methods (interview, observation, focus group)

Award

Erasmus Mundus scholarship 2008–2010
One of the 20 recipients selected from around the world, funded by European Commission

Press

"Mon smartphone me coache (My personal coach on smartphone)"
RTS Un (Swiss public television channel). 2012/11/8

Teaching Experiences

Introduction to Human Computer Interaction (CS-486), EPFL 2011–2014

Social activities

Board member of International Student Union, Trondheim, Norway 2009-2010

Language

Chinese (native), English (fluent)
French (basic), German (basic), Finnish (basic), Norwegian (basic)

Personal Interest

Meditation, painting, hiking, swimming